

<110> Boronat, Albert;  
Campos, Narciso;  
Rodriguez-Concepcion, Manuel;  
Rohmer, Michel;  
Seeman, Myriam;  
Valentin, Henry E.;  
Venkatesh, Tyamagondlu V.;  
Venkatramesh, Mylavarapu

<120> Methyl-D-Erythritol Phosphate Pathway Genes

<130> 16516.107/35-21(51897)US

<150> US 60/223,483

<151> 2000-08-07

<160> 85

<210> 1

<211> 2520

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (154)..(2376)

<400> 1

aaaaatcgtc aatccctctc aaactcttct caccactaat ttcttcctct ggaacattct 60

cttctctatt attttgattc ccttggcctc aacactgggtt tctcaattgc atgatcttgg 120

ctcgtcttca gttactttga ttcactgaga aaa atg gcg act gga gta ttg cca 174

Met Ala Thr Gly Val Leu Pro  
1 5

gct ccg gtt tct ggg atc aag ata ccg gat tcg aaa gtc ggg ttt ggt 222

Ala Pro Val Ser Gly Ile Lys Ile Pro Asp Ser Lys Val Gly Phe Gly  
10 15 20

aaa agc atg aat ctt gtg aga att tgt gat gtt agg agt cta aga tct 270

Lys Ser Met Asn Leu Val Arg Ile Cys Asp Val Arg Ser Leu Arg Ser  
25 30 35

gct agg aga aga gtt tcg gtt atc cgg aat tca aac caa ggc tct gat 318

Ala Arg Arg Arg Val Ser Val Ile Arg Asn Ser Asn Gln Gly Ser Asp  
40 45 50 55

tta gct gag ctt caa cct gca tcc gaa gga agc cct ctc tta gtg cca 366

Leu Ala Glu Leu Gln Pro Ala Ser Glu Gly Ser Pro Leu Leu Val Pro  
60 65 70

aga cag aaa tat tgt gaa tca ttg cat aag acg gtg aga agg aag act 414

Arg Gln Lys Tyr Cys Glu Ser Leu His Lys Thr Val Arg Arg Lys Thr  
75 80 85

cgt act gtt atg gtt gga aat gtc gcc ctt gga agc gaa cat ccg ata	462
Arg Thr Val Met Val Gly Asn Val Ala Leu Gly Ser Glu His Pro Ile	
90 95 100	
agg att caa acg atg act act tcg gat aca aaa gat att act gga act	510
Arg Ile Gln Thr Met Thr Thr Ser Asp Thr Lys Asp Ile Thr Gly Thr	
105 110 115	
gtt gat gag gtt atg aga ata gcg gat aaa gga gct gat att gta agg	558
Val Asp Glu Val Met Arg Ile Ala Asp Lys Gly Ala Asp Ile Val Arg	
120 125 130 135	
ata act gtt caa ggg aag aaa gag gcg gat gcg tgc ttt gaa ata aaa	606
Ile Thr Val Gln Gly Lys Lys Glu Ala Asp Ala Cys Phe Glu Ile Lys	
140 145 150	
gat aaa ctc gtt cag ctt aat tac aat ata ccg ctg gtt gca gat att	654
Asp Lys Leu Val Gln Leu Asn Tyr Asn Ile Pro Leu Val Ala Asp Ile	
155 160 165	
cat ttt gcc cct act gta gcc tta cga gtc gct gaa tgc ttt gac aag	702
His Phe Ala Pro Thr Val Ala Leu Arg Val Ala Glu Cys Phe Asp Lys	
170 175 180	
atc cgt gtc aac cca gga aat ttt gcg gac agg cgg gcc cag ttt gag	750
Ile Arg Val Asn Pro Gly Asn Phe Ala Asp Arg Arg Ala Gln Phe Glu	
185 190 195	
acg ata gat tat aca gaa gat gaa tat cag aaa gaa ctc cag cat atc	798
Thr Ile Asp Tyr Thr Glu Asp Glu Tyr Gln Lys Glu Leu Gln His Ile	
200 205 210 215	
gag cag gtc ttc act cct ttg gtt gag aaa tgc aaa aag tac ggg aga	846
Glu Gln Val Phe Thr Pro Leu Val Glu Lys Cys Lys Lys Tyr Gly Arg	
220 225 230	
gca atg cgt att ggg aca aat cat gga agt ctt tct gac cgt atc atg	894
Ala Met Arg Ile Gly Thr Asn His Gly Ser Leu Ser Asp Arg Ile Met	
235 240 245	
agc tat tac ggg gat tct ccc cga gga atg gtt gaa tct gcg ttt gag	942
Ser Tyr Tyr Gly Asp Ser Pro Arg Gly Met Val Glu Ser Ala Phe Glu	
250 255 260	
ttt gca aga ata tgt cgg aaa tta gac tat cac aac ttt gtt ttc tca	990
Phe Ala Arg Ile Cys Arg Lys Leu Asp Tyr His Asn Phe Val Phe Ser	
265 270 275	
atg aaa gcg agc aac cca gtg atc atg gtc cag gcg tac cgt tta ctt	1038
Met Lys Ala Ser Asn Pro Val Ile Met Val Gln Ala Tyr Arg Leu Leu	
280 285 290 295	
gtg gct gag atg tat gtt cat gga tgg gat tat cct ttg cat ttg gga	1086
Val Ala Glu Met Tyr Val His Gly Trp Asp Tyr Pro Leu His Leu Gly	
300 305 310	
gtt act gag gca gga gaa ggc gaa gat gga cgg atg aaa tct gcg att	1134

Val Thr Glu Ala Gly Glu Gly Glu Asp Gly Arg Met Lys Ser Ala Ile  
 315 320 325

gga att ggg acg ctt ctt cag gac ggg ctc ggt gac aca ata aga gtt 1182  
 Gly Ile Gly Thr Leu Leu Gln Asp Gly Leu Gly Asp Thr Ile Arg Val  
 330 335 340

tca ctg acg gag cca cca gaa gag gag ata gat ccc tgc agg cga ttg 1230  
 Ser Leu Thr Glu Pro Pro Glu Glu Glu Ile Asp Pro Cys Arg Arg Leu  
 345 350 355

gct aac ctc ggg aca aaa gct gcc aaa ctt caa caa ggc gca ccg ttt 1278  
 Ala Asn Leu Gly Thr Lys Ala Ala Lys Leu Gln Gln Gly Ala Pro Phe  
 360 365 370 375

gaa gaa aag cat agg cat tac ttt gat ttt cag cgt cgg acg ggt gat 1326  
 Glu Glu Lys His Arg His Tyr Phe Asp Phe Gln Arg Arg Thr Gly Asp  
 380 385 390

cta cct gta caa aaa gag gga gaa gag gtt gat tac aga aat gtc ctt 1374  
 Leu Pro Val Gln Lys Glu Gly Glu Glu Val Asp Tyr Arg Asn Val Leu  
 395 400 405

cac cgt gat ggt tct gtt ctg atg tcg att tct ctg gat caa cta aag 1422  
 His Arg Asp Gly Ser Val Leu Met Ser Ile Ser Leu Asp Gln Leu Lys  
 410 415 420

gca cct gaa ctc ctc tac aga tca ctc gct aca aag ctt gtc gtg ggt 1470  
 Ala Pro Glu Leu Leu Tyr Arg Ser Leu Ala Thr Lys Leu Val Val Gly  
 425 430 435

atg cca ttc aag gat ctg gca act gtt gat tca atc tta tta aga gag 1518  
 Met Pro Phe Lys Asp Leu Ala Thr Val Asp Ser Ile Leu Leu Arg Glu  
 440 445 450 455

cta ccg cct gta gat gat caa gtg gct cgt ttg gct cta aaa cgg ttg 1566  
 Leu Pro Pro Val Asp Asp Gln Val Ala Arg Leu Ala Leu Lys Arg Leu  
 460 465 470

att gat gtc agt atg gga gtt ata gca cct tta tca gag caa cta aca 1614  
 Ile Asp Val Ser Met Gly Val Ile Ala Pro Leu Ser Glu Gln Leu Thr  
 475 480 485

aag cca ttg ccc aat gcc atg gtt ctt gtc aac ctc aag gaa cta tct 1662  
 Lys Pro Leu Pro Asn Ala Met Val Leu Val Asn Leu Lys Glu Leu Ser  
 490 495 500

ggt ggc gct tac aag ctt ctc cct gaa ggt aca cgc ttg gtt gtc tct 1710  
 Gly Gly Ala Tyr Lys Leu Leu Pro Glu Gly Thr Arg Leu Val Val Ser  
 505 510 515

cta cga ggc gat gag cct tac gag gag ctt gaa ata ctc aaa aac att 1758  
 Leu Arg Gly Asp Glu Pro Tyr Glu Glu Leu Glu Ile Leu Lys Asn Ile  
 520 525 530 535

gat gct act atg att ctc cat gat gta cct ttc act gaa gac aaa gtt	1806
Asp Ala Thr Met Ile Leu His Asp Val Pro Phe Thr Glu Asp Lys Val	
540 545 550	
agc aga gta cat gca gct cgg agg cta ttc gag ttc tta tcc gag aat	1854
Ser Arg Val His Ala Ala Arg Arg Leu Phe Glu Phe Leu Ser Glu Asn	
555 560 565	
tca gtt aac ttt cct gtt att cat cac ata aac ttc cca acc gga atc	1902
Ser Val Asn Phe Pro Val Ile His His Ile Asn Phe Pro Thr Gly Ile	
570 575 580	
cac aga gac gaa ttg gtg att cat gca ggg aca tat gct gga ggc ctt	1950
His Arg Asp Glu Leu Val Ile His Ala Gly Thr Tyr Ala Gly Gly Leu	
585 590 595	
ctt gtg gat gga cta ggt gat ggc gta atg ctc gaa gca cct gac caa	1998
Leu Val Asp Gly Leu Gly Asp Gly Val Met Leu Glu Ala Pro Asp Gln	
600 605 610 615	
gat ttt gat ttt ctt agg aat act tcc ttc aac tta tta caa gga tgc	2046
Asp Phe Asp Phe Leu Arg Asn Thr Ser Phe Asn Leu Leu Gln Gly Cys	
620 625 630	
aga atg cgt aac act aag acg gaa tat gta tcg tgc ccg tct tgt gga	2094
Arg Met Arg Asn Thr Lys Thr Glu Tyr Val Ser Cys Pro Ser Cys Gly	
635 640 645	
aga acg ctt ttc gac ttg caa gaa atc agc gcc gag atc cga gaa aag	2142
Arg Thr Leu Phe Asp Leu Gln Glu Ile Ser Ala Glu Ile Arg Glu Lys	
650 655 660	
act tcc cat tta cct ggc gtt tcg atc gca atc atg gga tgc att gtg	2190
Thr Ser His Leu Pro Gly Val Ser Ile Ala Ile Met Gly Cys Ile Val	
665 670 675	
aat gga cca gga gaa atg gca gat gct gat ttc gga tat gta ggt ggt	2238
Asn Gly Pro Gly Glu Met Ala Asp Ala Asp Phe Gly Tyr Val Gly Gly	
680 685 690 695	
tct ccc gga aaa atc gac ctt tat gtc gga aag acg gtg gtg aag cgt	2286
Ser Pro Gly Lys Ile Asp Leu Tyr Val Gly Lys Thr Val Val Lys Arg	
700 705 710	
ggg ata gct atg acg gag gca aca gat gct ctg atc ggt ctg atc aaa	2334
Gly Ile Ala Met Thr Glu Ala Thr Asp Ala Leu Ile Gly Leu Ile Lys	
715 720 725	
gaa cat ggt cgt tgg gtc gac ccg ccc gtg gct gat gag tag	2376
Glu His Gly Arg Trp Val Asp Pro Pro Val Ala Asp Glu	
730 735 740	
atttcaaaac ggagaaagat gggtgggccca ttctttgaaa actgtgagag aagatatata	2436
tatttggtgtg tgtatatcat ctgtttgttg tgtattgcat catcattttg aacaaatgtc	2496
caaatctctt aagttgataa aagt	2520

<210> 2  
 <211> 33675  
 <212> DNA  
 <213> Oryza sativa

<220>  
 <221> CDS  
 <222> (6924)..(7019), (7163)..(7269), (7344)..(7444), (7525)..(7634),  
 <222> (7694)..(7813), (7923)..(8153), (8253)..(8369), (8515)..(8589),  
 <222> (9012)..(9071), (9163)..(9225), (9328)..(9472), (9589)..(9730),  
 <222> (9951)..(10028), (10134)..(10293), (10694)..(10798),  
 <222> (11028)..(11129)

<220>  
 <221> unsure  
 <222> (1..33675)  
 <223> unsure at all n locations

<400> 2

cttaaccctc gccgactgcc tggagattcg tgccgatcga tacacgtggc agcgcctaac 60  
 gcgtaacccc tccctcactt ggagattcgt gcaagcaact cgattaatgc attaatgctg 120  
 tcgcgtaggt ttccctacgg aagagctgag tttcgtaacg aaaaaaacg gccacgtttc 180  
 gcatcgagcc tactttaatt agcgtgggaa aataattcaa agtagcgacc tgtaccctgt 240  
 ggcaacctag cgcgcgcggc catggctctt gttccgctcg tgacagtgct cctgttcgcc 300  
 ggctcatgcc tcggatcagc gccgccgacg acatcgccgg cggcgctcggc ggcgctccacg 360  
 gcgacacgta cggtagtagt cgacggcatt acggccatct acaaacctcg gcgactcgct 420  
 gtcggacacc gcaacctcgc caggcaaggc gccaccggcg ggctgctccg gtacaccacg 480  
 aggcttccct acggcgtcac cgtcggccgc gccaccggcc ggtgctccga cggctacctc 540  
 atcatcgact tctcgggtga cgtcatcagt ttaatttctc tctctcttcc gtctgaaaaa 600  
 tggaagaaac aatattatat tacgttatat atatatgcgt ttttgtttcg gattaaattg 660  
 tggatatgat cgatcgatgt gcagctagag atcttggcct ccctctgctc aaccctgacc 720  
 tcgacgaggg cgcggacttc gccacggcg tcaacttcgc cgtcgccggc gccaccgcgc 780  
 tcaacacgac ggcgctcgcc gccaggcgga tcaccgtccc ccacaccaac agccccctcg 840  
 acgtgcagct cagatTTTTT ttgttttaga gaagggtatt ttttaccgg cctctacatc 900  
 caaccggata tatacggcta ttgaagtagg gaacttaacc ctgtaaacia tccatccata 960  
 gaggatatga acctaagacc ttgaggtact acttcaaccg gatataatcg tgcagctcag 1020  
 atggttcaag gaattcatga actccacaac tagttctcct caaggtgaac gaacaaactg 1080

aaacgcattt	cagcttaatt	tcgaccggtg	cctgatcagt	gccagtcagc	aatgctgtat	1140
ctcacaaata	attaagctaa	tgtacagctt	ttcagtgcta	gaatgacttt	catatagaga	1200
aatcttgtgt	tatatatata	tacttttttc	tgaaagaaaa	aagtcttttt	gtgtgagcat	1260
tgcattgcag	agatccgtga	aaagctgtcg	aagtcactgg	ttatgctggg	agagatcgga	1320
ggaaacgact	acaactacgc	cttcctccag	acctggccga	tggacggtgg	atacagcctc	1380
ggcaacgtca	cacgcatgat	cgaaagcggt	gccacggccg	tcgatcttgt	accggaagtc	1440
gtgcagtgca	tagccagcgc	agccaaggta	cacaccattc	ttttccatta	atttttggga	1500
ccttattttt	aaaataataa	tcctggctac	aaagtaatta	attaagaact	aaattaattt	1560
ttgtgggttt	tgtgacacag	gaggtgctcg	acatgggccc	gacgcgggtg	gtgatcccg	1620
gcaacctccc	gctgggttgc	gtgccgagct	acatgagcgc	ggtgaacgcg	acggaccggg	1680
cggcgtacga	cgcccgcgga	tgcctcgctc	cgctcaacct	cttcgcggcg	ctgcacaacg	1740
cgtggctgcg	ccgcgccgtc	ggggagctgc	ggcgcgcgta	ccggggcgcc	gcggtggctc	1800
cgtacgcgga	ctactccgcc	gcgtacgccg	cgacgctgga	cggggcagcg	gcgctcggct	1860
tcgacgagcg	gcgcgtgttc	agggcgctgt	gcggcaaggg	cggcgggggc	gcgtacgggt	1920
tcgacgtgcg	cgcgatgtgc	ggcgcgccgg	ggacggcggc	gtgcgcggac	ccggggaggt	1980
acgtgagctg	ggacggcgtc	cacctgacgc	agcgcgcgta	cggcgtcatg	gccgagctgc	2040
tgttccgccg	tggcctcggt	caccgcctc	cgataaattt	cacgaacagc	gcgcgcgcgt	2100
gaggcggtgt	tgcattggctt	gcgcgttttt	tctgatcaaa	actactcaag	tttgagccgt	2160
tttgatttat	aaataaaacc	atatgcgatt	ttgctaaacg	tttgctgcgt	gatttctctt	2220
cggaagaaaa	aatctcacc	gagtgatgca	taggcgggtc	caaccatatg	tgccctgacc	2280
tttctctgct	tccttcgcgt	cgtgcactga	caacctcaca	gtatgttttt	gggatgggcg	2340
cttgcgggccc	aactcaatct	gtaatacatt	gggctgtcgt	attgggtttg	ttggacttca	2400
tagactggat	cggagaaagt	tggttaattg	actttttcat	ttttgctata	aatgatttaa	2460
ttaaacagtc	taggataatt	actgtagact	ctaataatat	tgtttgggta	agtattatta	2520
tacattcctg	tatttgacac	tctaagagca	tggccaagag	ttgcctgaaa	gtctcttcct	2580
aaatctgcct	ttcattctct	aatgagaatt	taaggattaa	aaatatactt	attttcaata	2640
gacagcataa	atttaattcc	ctagaataaa	aaaatgcccc	cctaacaaca	gaaattagat	2700
tcctctaccc	gcacctcatc	agatcgctcg	atttaagatc	acgccatctg	acaccgccct	2760

cccgctcgct	cttctctagt	gtgggagtct	cgcgctcaag	agacggaaat	cggaacaag	2820
aatgattcct	agcttagcga	gaatgaagg	gaagacatat	gtcacaccta	caccacata	2880
agtatgccct	agcacaagg	atgaaaacg	atcgaaaacg	gatggaaact	agctttatca	2940
tattcgTTTT	cattttttt	tcggaatcg	attcgaaatc	gaaaactcg	atacggaaat	3000
aaaattgaat	attatcgaat	acagatacgg	agcgaatata	agatggaacg	aatacagtag	3060
cgaatattta	ccggtatata	aaaaaccct	caaattgagt	ttcttgatta	agaaagagat	3120
atcgcttatt	atTTtagtta	aatatctcca	acatttatat	cgtcaatttt	atagacggtt	3180
ccacaatcgt	atgtgaaaat	cgattttcat	ggttggttct	ctaagagatc	catatgcaaa	3240
tatgattatc	atTTtctatt	ccaagacctt	ttactagatg	tataacttat	ttaccattgc	3300
ataaattgga	gatgttattt	atTTtacttc	acatcttcga	aacttgtaat	gtatgtatta	3360
tactttaaat	gctttcaagt	acaaatgtta	taaactacaa	agtggtagat	ccggttgagc	3420
tctacaactt	tgatatggaa	cacatctcca	tcagatgtcg	tttgaattgt	agatctgaga	3480
ttttgtaaaa	tttaatatgg	tatattataa	tgaatattta	gacccttaa	tgaccttaa	3540
taataaaata	gtcaataata	aagttgtaga	tctcatcgag	ctctataatg	ttgatatgaa	3600
gtttgtcttc	atctgattcc	gtatgaaaa	gttatgtata	tatacatggt	tttttataaa	3660
atTTgtcaa	tatctgcgga	tatccgaaaa	aaatttcgga	tagtttttaa	ccgtttttcg	3720
attccgatgg	atagtatcct	tactgtattc	gttttcgttt	ccgagaaaa	atatccaaat	3780
tcgtttccga	atccgagaat	ttttggataa	ttccgacaga	aactatccga	atccgaaaa	3840
tggttcggac	ggacggaaac	tatccaaacc	agtttcatcc	cggctagcac	gcattttaa	3900
tcacatgagg	ttgcacattt	atctgaggta	aaaagattgg	aaacggttac	tggttcgtca	3960
agaattttcc	gtatttatca	gtataactat	tcaatgacga	catcaacata	acagaaaatt	4020
aaaacaacat	gagtcgattt	tatatataac	tagaaacgaa	aacagtataa	ctgttacgaa	4080
aacactagat	tgatgggtcg	aaaatttcca	ccacggtttt	tatgcctacc	tttcaacgct	4140
cccaaagttc	ccacgacca	aaacatgtgt	gggagaactt	ccgcccacat	ggagacggtt	4200
gtctcagga	aacgtgccat	ctgctttgct	ccaggtcaac	acatgtggtg	tgactgaact	4260
ggccatcgtc	tcaatattgt	catctaccg	tcataccatg	ccaccggacc	agaaggtgat	4320
tatggtcttc	ggcgccgctc	gcgcgcggat	gccttgctcc	acaacaagtc	agccgctcaa	4380
accacactcc	cctttggcat	tgaacatgag	gtttgacgac	gatgtgtgtg	tatgtttggg	4440
caggtagctt	tgTTtcaagc	tgactagct	aattaagatc	gatctccttg	tcaaagtcac	4500

gatcaaacat	cgaaagtaca	tgcatggaag	aatgttgaa	atgtaatgaa	ctaaatgatg	4560
tccttttctc	cccttattaa	acaacatcaa	gtttctttta	tttctaaaga	atgttaatat	4620
cctttttatt	tcttcaataa	atagtactgc	actccctatg	gtttttgttg	tttagcatct	4680
tgactttcgg	gcatacgttt	tatgattttat	cttattacaa	aatataatta	tcattttattt	4740
tatcattaca	aatactttta	aaataacatt	atcagctgat	tttgaattaa	aactaaaatt	4800
acaccttaat	tacaatatac	ttcacatagc	aattataata	taactatata	caacttacac	4860
tataagttat	gttcaaaaata	tttttcctac	aaaaactatc	accagattct	tagacagtcc	4920
cattccacca	cctcagctgc	cgtgaaagaa	ctttgggtct	taaataagtc	caaatttatc	4980
tttttgtttt	ctcaataaaa	tattcgaatt	atccaacaaa	tcaaggaaaa	aacatccttc	5040
gatgacccat	gaatattcgt	gaagtttctc	ctctagccag	taacaatacg	gaacaatcag	5100
acaattttat	ctggctcaag	caccatctct	cgcaccagat	taaactattt	ttttttcatg	5160
gtacaataca	atcccatgcc	ggccacgaaa	aacaaatggc	agaaataata	aacgaacaaa	5220
acagcctctc	tccatcgtga	actaataaaa	aataaaataa	aaacaaaaca	aatgataat	5280
ggaattacga	agcgcacggg	aaaacgacgg	gcacgattaa	atcatggcgg	ggagagcccg	5340
gaacccccact	tccacacctc	caacccccacg	ccgtcagcct	ccccctccca	tgacccccgt	5400
ccaccaaacac	ctcatctctt	ggacccccaca	cgcagccact	gccacaggca	acgcggtgct	5460
cgtgcaccga	gtccacacga	cgcgcgcgcg	ggtgcggggg	cgcgggcctc	tggggataaa	5520
tgggctaatac	cggtagaaaag	cccaccactc	gctcgccagt	tcgtcgctct	cttcgccgag	5580
ctcgcgagct	ctcgcaactct	gtctccatcc	ccgcacgcga	tcgcctcgcc	gctgctgatc	5640
tcgtcgcggt	cgcgggaggg	gagctacgag	gttggggagc	cttatctcta	cttcctgaga	5700
tttctagtag	ctttgtgtat	gtgtgtgtgt	ttgtgtgttg	gggggacgcc	gatcgggttg	5760
atcctcctgt	ggtggttggt	tgggcgcaat	tcgtgcttgg	tttatttgct	ggaattctag	5820
cgggggagct	ggcgttgctg	gtgctaattg	ctgcggggga	gctgctggaa	ttcgtgcttc	5880
tgcttgggaa	ttagaaggtt	tgggttttta	tgattcagag	ggctgtagag	ctcttgagat	5940
tggctgcgaa	aattcgggat	ttgatcaact	tagagagcat	tatctttgga	ttaggagggg	6000
tttttcttaa	tttttcttag	ttttttttga	gctatcaaga	gttcatgcc	tcttatttct	6060
ccctttgttc	ttagccggaa	ggatacacga	atcagttttt	tttttttaaa	aaaaatattt	6120
atctcaattt	tctgcaagca	tgttcaattt	ctaagtggaa	atgctattta	aaagaccagg	6180



cttattgatt ggtgctatac tttgattttc tttggaattg tagtagaagc atcagtttct	6240
tcatgctgtc ctaccaacct ctcttattat tagcaaagta aagttattaa atttgctaatt	6300
tgttgatatg tcagtatttt gtacgaattg tgaaatagtt aattttcaat aactacacac	6360
catggttgtc ctggtgttgg actggaagca ataaggggaat attccatttc tgtccattaa	6420
aaccacaaaa gatgaccctg tgctcatctc taccattgcc atgcacctgt ttgtaggatt	6480
gcctaaccce gaagttgggtg ctctgagata gccatggcca ccggagtggc accagcgccg	6540
ctcccatatg tcaggggccg tgatgggtggc atcggttca cgaggagcgt cgactttgct	6600
aagatcttgt cgggttctgc tactctaagg gtgggtcat caagaggcag ggtgcttgtg	6660
gccaagagct caagtaccgg ttctgatacc atggagctcg agccatcttc agaaggaagc	6720
ccacttttag gtataactcg ccggctgttg ttcacctgc atgtatatc gtgttagttg	6780
ttcttagtgc ttttaactga atgaacattt tttctgtaa gaatctgaca gcatgtcttt	6840
tgcccttttg ttattcttta gttcccaggc aaaagtattg tgaatctata tatgagacaa	6900
ggaggagaaa aaccgcact gtg atg gtt ggg aat gtg cca ctt ggc agt gat	6953
Met Val Gly Asn Val Pro Leu Gly Ser Asp	
1 5 10	
cat ccc att agg att cag act atg acc acc tcg gat acc aag gat gtt	7001
His Pro Ile Arg Ile Gln Thr Met Thr Thr Ser Asp Thr Lys Asp Val	
15 20 25	
gct aaa acc gta gag gag gtacactcct atttgaagtt ctatgtttta	7049
Ala Lys Thr Val Glu Glu	
30	
gtttttaatt ctatgcttga ataattgaat gctgggcatg cattaatcat gtgttctttt	7109
agatgttcta tgtttcatga ctagtgaaat aacgaagtat agcactggtc cag gtt	7165
Val	
atg agg ata gca gat aaa ggg gct gat ttt gtt aga ata aca gtc cag	7213
Met Arg Ile Ala Asp Lys Gly Ala Asp Phe Val Arg Ile Thr Val Gln	
35 40 45	
ggt aga aag gaa gct gat gcc tgc ttt gag att aag aac act ctt gtt	7261
Gly Arg Lys Glu Ala Asp Ala Cys Phe Glu Ile Lys Asn Thr Leu Val	
50 55 60 65	
cag aag aa gtaagagtca tcatttttcc agattcagtg agttttcatg	7309
Gln Lys Asn	
aatgaattct catcttgctt ttgcatttca acag t tac aac atc ccc cta gtg	7362
Tyr Asn Ile Pro Leu Val	
70	

gct gat att cat ttt gcc ccg aca gtt gct tta aga gtg gct gaa tgc	7410
Ala Asp Ile His Phe Ala Pro Thr Val Ala Leu Arg Val Ala Glu Cys	
75 80 85 90	
ttt gac aaa att cgt gtc aac cca ggg aat ttt g gtgagtgaaa	7454
Phe Asp Lys Ile Arg Val Asn Pro Gly Asn Phe	
95 100	
taatgatgtg tatcatttta gtgtcaatat cttatcaact ctgtgcatat gctgagaact	7514
ctacttgacag ct gat cgc cgt gcc caa ttt gag cag ctt gaa tat act	7562
Ala Asp Arg Arg Ala Gln Phe Glu Gln Leu Glu Tyr Thr	
105 110	
gaa gat gat tat caa aaa gag ctt gag cat atc gag aag gtt cca aat	7610
Glu Asp Asp Tyr Gln Lys Glu Leu Glu His Ile Glu Lys Val Pro Asn	
115 120 125 130	
atc tca ctc ttt agt gtt aat tta gtcagtaaga atgtgcagta tgtttcctta	7664
Ile Ser Leu Phe Ser Val Asn Leu	
135	
cttgcatagc cacttccata tcatttcag gtc ttc tcc ccg ttg gtt gag aaa	7717
Val Phe Ser Pro Leu Val Glu Lys	
140 145	
tgc aag cag tat gga aga gca atg cgt ata gga aca aat cat gga agt	7765
Cys Lys Gln Tyr Gly Arg Ala Met Arg Ile Gly Thr Asn His Gly Ser	
150 155 160	
ctg tct gac cgc ata atg agt tac tat ggt gat tct cca cgc gga atg	7813
Leu Ser Asp Arg Ile Met Ser Tyr Tyr Gly Asp Ser Pro Arg Gly Met	
165 170 175	
gtattatttc ctttctgggg atttcattca aataactttt cgtttcatgg atgtcttcaa	7873
ttaatgatcg ttttgataga tgaatgacat gttctacaaa taatttcag gtt gag tct	7931
Val Glu Ser	
180	
gct ttg gaa ttt gcc agg atc tgt cgg aag ctg gac ttc cat aac ttt	7979
Ala Leu Glu Phe Ala Arg Ile Cys Arg Lys Leu Asp Phe His Asn Phe	
185 190 195	
gtg ttt tca atg aaa gca agt aac cct gtt atc atg gtc caa gca tat	8027
Val Phe Ser Met Lys Ala Ser Asn Pro Val Ile Met Val Gln Ala Tyr	
200 205 210	
cgc ttg ctt gta gca gaa atg tat aac cta ggg tgg gat tat cct ttg	8075
Arg Leu Leu Val Ala Glu Met Tyr Asn Leu Gly Trp Asp Tyr Pro Leu	
215 220 225	
cac ttg gga gtt aca gaa gct gga gag ggt gaa gat ggg agg atg aag	8123
His Leu Gly Val Thr Glu Ala Gly Glu Gly Glu Asp Gly Arg Met Lys	
230 235 240 245	

tct gcc att ggc att gga aca ctt ctg atg gtaattgcat ttttactttg	8173
Ser Ala Ile Gly Ile Gly Thr Leu Leu Met	
250 255	
tgtattatat tgcatatatc atatctttcc atctgcaaag ggtaagcatg ccttatgtct	8233
tcctttttgtt gtctttacag gat ggc ttg ggc gat aca atc cgt gtc tcc ctc	8285
Asp Gly Leu Gly Asp Thr Ile Arg Val Ser Leu	
260 265	
acg gaa cca cct gaa gaa gag att gat cct tgc cgg aga ttg gca aat	8333
Thr Glu Pro Pro Glu Glu Glu Ile Asp Pro Cys Arg Arg Leu Ala Asn	
270 275 280	
ctt ggc aca cat gcc gca gac ctt caa ata gga gtg gtaacgattt	8379
Leu Gly Thr His Ala Ala Asp Leu Gln Ile Gly Val	
285 290	
attacctttc tctagtttta cactttttctc ttgttttagct gccaatgccca cacattaatt	8439
ttgactatatt ttagtagtgt tttgttctat ttgttctttt aagaatttct atttatatac	8499
attatatgtt ctcag gct cct ttt gaa gaa aag cac agg cgc tat ttt gat	8550
Ala Pro Phe Glu Glu Lys His Arg Arg Tyr Phe Asp	
295 300 305	
ttc cag cgt aga agt ggt cag ttg cct tta caa aag gag gttagttcaa	8599
Phe Gln Arg Arg Ser Gly Gln Leu Pro Leu Gln Lys Glu	
310 315	
aataactcct atagtccata gttatcataa aaacaatagt gctagatttc ttattagttg	8659
cacttatgac aggggtgagga agtagactac agaggggtct tgcaccgtga tggctctgtt	8719
ttgatgtcag tttccttgga tcagttgaag gtaactcaca tatttgttac ccttttgtgc	8779
aatgtgttga tcttgtgtaa ctttaccaaa atatatttca agacaatagt ctatttttga	8839
atatacaatt ctacaacatg atattttcag tagccatgtt ccatgcattc tatgcatagt	8899
tcatagtaca tagtgagaat agcaatagca aaaagaaggc attgattttt ttctatctga	8959
atcaaatacaa ttgatgcatt ttgtaatgat ggaaggctct cttatttttc ag gct cct	9017
Ala Pro	
320	
gag ctc ctt tat agg tct ctt gct gca aag ctt gtg gtt ggc atg cct	9065
Glu Leu Leu Tyr Arg Ser Leu Ala Ala Lys Leu Val Val Gly Met Pro	
325 330 335	
ttc aag gtctgacct tatagctgta cattctagca aacaactaaa ctttattgggt	9121
Phe Lys	
acttcagtct aaactgatgt taatttttct atgaatatca g gat ctg gca act gta	9177
Asp Leu Ala Thr Val	
340	

gat tct att ctt ttg aag gag ctc cca cct gta gaa gat gct caa gct	9225
Asp Ser Ile Leu Leu Lys Glu Leu Pro Pro Val Glu Asp Ala Gln Ala	
345 350 355 360	
gtgagttcct tcaacattat ttgttctttt cacaaatcac aagcttatat taacattcta	9285
ttccttttaaa atttttgtgt tgaaatctgt aaaatggtag ag agg ctt gca ctc	9339
Arg Leu Ala Leu	
aaa aga tta gtt gac atc agc atg ggt gtg ttg act ccc tta tca gag	9387
Lys Arg Leu Val Asp Ile Ser Met Gly Val Leu Thr Pro Leu Ser Glu	
365 370 375 380	
caa ctg aca aag cca ctc cca cat gca att gct ctt gtc aat gtg gat	9435
Gln Leu Thr Lys Pro Leu Pro His Ala Ile Ala Leu Val Asn Val Asp	
385 390 395	
gaa ctg tca agc ggt gca cac aaa ctt ttg cca gaa g gtagacattt	9482
Glu Leu Ser Ser Gly Ala His Lys Leu Leu Pro Glu	
400 405	
gaatttgata atgatctttg ttgttttggtg aattgtgttt atgtcatttt ctgtatttta	9542
acattttgct tagtctgttt tattgatgaa tctttttttt atgtag gc act aga	9596
Gly Thr Arg	
410	
ttg gct gtc acc ctt cgt gga gat gaa tca tat gaa cag cta gat ctt	9644
Leu Ala Val Thr Leu Arg Gly Asp Glu Ser Tyr Glu Gln Leu Asp Leu	
415 420 425	
ctt aag ggt gtt gat gat ata aca atg tta ctg cac agt gtt cct tat	9692
Leu Lys Gly Val Asp Asp Ile Thr Met Leu Leu His Ser Val Pro Tyr	
430 435 440	
ggt gaa gag aag act ggc aga gta cac gct gct agg ag gtaagtgaac	9740
Gly Glu Glu Lys Thr Gly Arg Val His Ala Ala Arg Arg	
445 450 455	
acagtaggcc agttaatacc actccctcca ttattaccat ttgttgggat gaaccgatag	9800
tcaattctaa gttacacatt aagcatgaaa aatgaaaatg gatttgactc tgcagaaaac	9860
tgacatacag accaatgttt ccacctgggt ttccattggt ctgtacttct ctttacctaa	9920
aattttatatt tttttaataa tgttttgcag g tta ttt gag tac tta gaa acc	9972
Leu Phe Glu Tyr Leu Glu Thr	
460	
aac ggt ttg aac ttc cct gta atc cat cac ata gaa ttc ccc aaa agc	10020
Asn Gly Leu Asn Phe Pro Val Ile His His Ile Glu Phe Pro Lys Ser	
465 470 475	
gtg aac ag gtactatgaa gtgcttatta agagatgcat tgaccgcccc	10068
Val Asn Arg	
480	

tccttacccc ttgaaattac tgtaccttta ttctcttggtg cttatttgag ttaaattata	10128
tgacag a gat gac ctt gtt att ggt gct ggg gca aat gtt ggt gct ctt Asp Asp Leu Val Ile Gly Ala Gly Ala Asn Val Gly Ala Leu 485 490 495	10176
cta gtt gat ggt ctt ggt gat ggt gta ctt ctt gaa gct gct gac cag Leu Val Asp Gly Leu Gly Asp Gly Val Leu Leu Glu Ala Ala Asp Gln 500 505 510	10224
gaa ttt gag ttt ttg agg gac aca tcc ttc aac ttg tta cag ggc tgc Glu Phe Glu Phe Leu Arg Asp Thr Ser Phe Asn Leu Leu Gln Gly Cys 515 520 525	10272
agg atg cgc aac aca aaa acg gtaagctgat gaattcttct ctgtagact Arg Met Arg Asn Thr Lys Thr 530 535	10323
gtagatccca tgaacaacgt caacctttta ctcgtgagat atcatgaaga agtgcaaaat	10383
tgcaactttta acagtaaagt aaccttatag cctaccgaag aggataaata actttaggca	10443
attctctctt gtgaagcaga acattctttt ggcgatttct gaccgttaat taatgctgca	10503
ggaatatgtc tcttgctcct cttgtgggcg gacactcttt gacctccaag aagtcagtgc	10563
tcagattaga gagaagacct ctcatctgcc aggcgtctct gtaaactctc ttacagacct	10623
tctgcctccc ttgttttcaa tcgcatatta gctagcctga tggctaataca tgtctacatt	10683
tgccctggcag att gct atc atg ggt tgc att gtc aat ggg cca ggg gag Ile Ala Ile Met Gly Cys Ile Val Asn Gly Pro Gly Glu 540 545	10732
atg gcc gat gct gat ttc gga tac gtt gga ggt gct cct ggg aag atc Met Ala Asp Ala Asp Phe Gly Tyr Val Gly Gly Ala Pro Gly Lys Ile 550 555 560	10780
gac ctt tat gtt ggc aag gtaacctttt cctatacttg tggaagttga Asp Leu Tyr Val Gly Lys 565 570	10828
atcatatcaa atggaataat ggaaatcacg gtatatcggt gaacatagct gcaagtcaat	10888
atttgtagat gatcatgcaa acacaatcaa cagtagggat gttaactgca tggcatatat	10948
atgctctttg agctgaaaca aaaacttaga gctgccattt tccttccatt aacacaagtt	11008
ctacttgttt tgggtgcag acc gtc gtg caa cgg ggc att gca atg gag ggg Thr Val Val Gln Arg Gly Ile Ala Met Glu Gly 575 580	11060
gcc act gac gcc ttg att cag tta atc aag gac cat ggc cgt tgg gtg Ala Thr Asp Ala Leu Ile Gln Leu Ile Lys Asp His Gly Arg Trp Val 585 590 595	11108

gat cct cct gtt gag gag tag gccgtagcat gtagttcata tatgtactcc 11159  
 Asp Pro Pro Val Glu Glu  
 600

tccataaaca atgttgtagc tgaggcacat tgtattgtat ccacggagta cataaataca 11219  
 cgttctgtac atcagtttag aaataaagta ggaatagggg tggctgcaac tttgtaacac 11279  
 cctcgtgaag catcggaaca tccaaattag aagcgtcctg aaatcagtga aaaagaattg 11339  
 atactgctat tttttgtacc aattgaaaaa aaaaaggaat acatgatatg actaaatcat 11399  
 gggttacatc ttcgtcaaaa aatgtcacag cttacattat tttcactact tgcaaatacc 11459  
 agacgatcta ctggtgcggg aacttgacgg gtgcaggaga cgcgaagccc ttgtggtaga 11519  
 gaagctcggc catgacgctg tacgcgcgct gagtcagggtg gacgccacgc catcccagct 11579  
 gatctgctcc atctcgaagt tgtacttccc gccgcagcac gccttggtca gcgccacgcc 11639  
 gtgcaacccc gtgtcgcgcg cgccctccag catccgcacg tacgcgccgg agtagtcggc 11699  
 gtacgcgatc gtggcctccg gttatgaccg cctcagctcc cggatcccct gctgcagcag 11759  
 cacgttggtg atctgcgcga acaggttgag acccagagg caccggttcc cgtcgtacgc 11819  
 cgcgcgctcc gtctcgtcca ccgccgccag gtagctcggc gcgcaaccca gcgggaagtt 11879  
 gcccgggatc accaccgcg tcgcgctcat ctgcagcacc tccctcgccg cgctcaccac 11939  
 gcgaccgcac cacctctggt acgagcacca ccgactccac cacgccggtc atcatgcgcc 11999  
 cgacgtccgc gcggcgctac gacctctgt tcggggcctg ttcggggcga tctctccac 12059  
 catcaccagc aagctcgcgc cagcttctct tgctcgagaa ttttcagaat atgccaccga 12119  
 atatgcaccg ttttcaggat agaccactca attcgacta ctttcataat atggcatttg 12179  
 gacgcgatat tttcttcggt ccgtgacact ctcatccttc caccgtcagc gccagtaatt 12239  
 ccgttcgcac accaacagct ctctctgagc gtccagctcc agtgggggag ttgttggtgc 12299  
 gcggcgcggt aacaccgatc ctgcgcaggg ccgccgcgtc gagggcggtg gcgccggtga 12359  
 cggcgaaagt tgacaccgta ggagaagtcg gcgcctttgt cgatgtacgg gttgagcagc 12419  
 ggcagcccta ggtcgttggc gaggtagtcg atcatgaggt acccgtcgtc ggagcactgc 12479  
 cccgtggcgc tgccgatggc cgcgccgtac gtagggaggc gccacggtgt gctccatcaa 12539  
 ggcgaggaag ttgccggtgt ccgagatgga gtccccgaag ttgtagatgt ccgtgatgcc 12599  
 gtccaccacc gccccttcg ccgccgatga caacgacgac gacgcggcct tccccggagc 12659  
 cggccttgcc tggcaagtgc cgacgaggag cagcgccaag aacgcgacga ggattggatg 12719  
 aaccggccta ctgcctatgg cgctcggtgc aagtgcgaagt ggggtgcgacg cagcagttgt 12779



gtccaacatg	tccgctgtgc	ttgtttgctg	ttgcaacacg	aacataatta	attactcaat	14519
tggttgcatt	attcatgcgc	aaaaaatgtt	accgctaatt	aatattagct	agaactagat	14579
gagagaacgt	acgacccctt	tcatctatat	acaataatca	tgaatttggt	gagaaagcat	14639
gtttggtatg	gtgttgaggt	tgtggctgtc	atgcaccaaa	gctctaactc	cagtgcctat	14699
agaatttaac	tacacaaaca	tggatacgct	ttttctagaa	attctattag	gttatgattt	14759
tgcgcttggt	gtccatgaat	ttgttgagca	tgtgttaagg	gacacttcac	agtgcacact	14819
catgggtgaa	tgcgtgtgca	tttgccatgt	ctattattaa	ggcgagaaac	atgaatctgt	14879
gtgctaattg	cacaagaaat	gtggaaagtt	tttttttaaa	agaaaatact	tagctaggga	14939
tgttcctttc	ttcctcaaat	atcatgtaaa	tataggtagt	aacattatgc	aaagttcaaa	14999
tcgtaatggc	caccttgtcc	atgttgggca	ccagctcctg	cagcttcctg	agcttctcgc	15059
taattctcgt	cctccgttcc	tacggacgcg	catcgatcac	accgacgtac	atgctcatgt	15119
gtcaagatct	gaagagaaag	caaaaagcaaa	tatagaggcg	ttttgatcat	gatattgcgt	15179
acgtaccctc	tccgcgatgc	tcctgggggtg	cgtcgcgcag	ccgcgcttgg	cccgcacttt	15239
gaacggcacc	tggatcatgt	gcagctgcag	gtacctgtcc	atgccggcca	tctccagcgc	15299
cgacgtgctc	gccatgccgc	cgaactgccc	ccatttccaa	cacgcccag	aaatcagaac	15359
acatcgcgat	atatatatat	atatatatat	atatatatat	atatatatat	atatatatat	15419
cacaaacaca	gcaaagctag	ctactacttc	ttcctctggt	ttacattaat	tattataagt	15479
tgttttgagt	tttgaataga	ttcatacatg	tataaatgta	tgtgtttcat	acatgtgtcc	15539
aaattcttat	gaatgttagt	aaatataaac	aagggatgaa	gagatcaaga	agccttgtag	15599
tgtacaatga	ttcaatgaag	gtagccctag	caatcaaatt	tgccgagcaa	tctttacctg	15659
ggactcgtac	ccgccgaggg	tggagatgat	gtccctggac	tcctcccacg	gcccgcagat	15719
ggagaacccg	ccgccgccgc	tgctgcgcgc	ggcgagagaag	gtgcggggca	cggaggcctc	15779
ggcgccggcg	cggtcgggga	aggcgccgtc	ctccgcgatg	tgcgagaggt	gcggcgcccc	15839
ggccgtgaag	ctcagctggg	acttcatctt	cctcccgcgc	ctgctgctgc	cgctgccgct	15899
gccggccatg	gaaggggtgt	ggtgggcttc	ggctccgctg	cctccccgc	cttttgagcc	15959
tggaaagcct	gcttagttta	ttgccaagta	gcaagcacgg	aaattaacta	atgatcgcta	16019
attagttaaa	ttaactgtgt	gtgtgagaga	aagagctact	gttaccctaa	cgctagttga	16079
aaactgccaa	gtgtgacaag	taaacaatag	tttacgggtat	tagcataccg	ttagagctag	16139
ctctataggt	acacgtgttg	agcaataagt	ttaacctaga	tgtgatggga	tgttcaaact	16199





gtgctaata	cggaataatt	aggcttaaaa	gattcgtctc	gcagtttcca	tgcaagttat	17939
gaaattat	ttttcattcg	tatctgaaaa	acccttcgga	catccggtca	aacatccgat	17999
atgacacca	aaatgtttct	tttcgcaaac	taaacaggcc	cttagcaaaa	tggttggtta	18059
tcaactttta	aaatatgttg	acagtgtctg	tgacgacttc	atgacggtcc	tctttaaagg	18119
tgcttatata	gtgatagggt	gtgcgtgtat	gttcagagcg	ttgagtatgc	atgtgtatat	18179
atgcatgttt	gtgtctgtac	tgtgttaaaa	aagaaaatcc	caagatctag	cctaaaattt	18239
tcattaaaaa	cattgaaatt	ttggccccac	gattttttta	ttccacaatg	taaatttcta	18299
gtcaaattgc	tgcaaatgac	gcgaaaatta	ttttctgacc	agtgaactga	catgcacaca	18359
ttacactata	tttattttat	atttattttg	aacgtacct	cgactacttc	caggggatcg	18419
atcttattct	cctcaaatta	ataagaacaa	gtactctctc	catttcaaaa	tacaacaacc	18479
taagaatatg	gataattttc	ttcattgaat	cggatggttt	cttcggtttt	tttgtactac	18539
gatgcgaaca	gatggtatat	tgaagcctac	cggacacgct	agcacgtgca	tgccgcgtgc	18599
cggcccggtc	atatgagcaa	gcctcgcacg	ctgacataga	cgcagccaag	agagaaagca	18659
aacgccaaat	caagaagccg	agcaatcacg	catgccatct	caacgcaccg	taggtcacta	18719
tcttttagcg	ggcaagaccg	tgacgtcacc	gtcaggccat	cagcagagga	gctgaacctg	18779
gacaaaccgg	ggggcccacc	ccgcaggcca	agttgcggcg	acacacacgt	gggtccccgcc	18839
ttacattaag	gcaagtggcg	ccctaattaa	tccattgatc	aaaaattaat	taatccacaa	18899
attaatcaaa	tgccctcacc	tttttctttt	tgccctggct	agggttcgag	gcactaagat	18959
ccactggtaa	tttaattgtg	cttgctgtct	tgatacta	taattgatca	tatatgcgca	19019
agttgggtcta	tctagagcag	aatctagagt	gcaactggct	gccgcattga	aagaaatgct	19079
gctacatggg	ctccactgaa	agacatttga	ctctttttaa	ctttactcga	ggctattcct	19139
acctgatca	aagtataatt	actaaattta	gtactgggtg	agtacttata	tgtggatttc	19199
gacatttcta	ctggtactat	ttttatcctt	accaattggt	gtatacagg	tgctcgggtca	19259
aaaggccatt	ttagatgttg	gtatatatgt	agtggtgaaa	tttaattata	cataactcta	19319
tgttcatatt	gatctgcatt	tcaaaaagat	attgacacac	ttattcctaa	tttttgaata	19379
aatgatattt	tgaagttttc	attaaagggt	tattatctct	gtatgctcta	aaacgttgaa	19439
tatttgtgac	gcagaattaa	tttaatactc	atggaataaa	taatgatgg	gcataatttt	19499
gcaatgattt	tcatcaaatg	aggtgcatat	aggtatcctt	tatatgaaat	gagaataactt	19559
gccaaaaaca	tttttaaaaa	gagcttggtt	tagctagcta	ggttgggtgaa	tggtgatact	19619

aattaatcaa atgtacatat ttgtgcaaat cctggaagat gaatgcatgg ttttctagtc 19679  
 ttattatgaa caaattaaat tagaaaaaaa aacatctatc tctttgctct ctccactata 19739  
 gcttcaaatt gttttttttc cccatgtcta ctattgtagt gaagaatgga ttgtcatgcg 19799  
 caatgacttt gcaactgaaa ataatggatc aaatgagaga gagggacacc aggtgcaagt 19859  
 ggcaaaaaaa ctaagccatt tatagcaagt tgcaatagaa aataagacaa tctagagaca 19919  
 ctcgattata aaaagcgta cgtaaaaaga taaaagcggg gtattcaaaa ccctagaccc 19979  
 cacatttcac tatcgatgat accctacttg agaaaacccg cctcctgtgt agcccatagt 20039  
 tttccatcgt ctttcttaca cgccgagcca aatttgtgca ctctctgtaa taacatatgc 20099  
 cttaaaaaact tgaactcata ttacattatc acgaaaacaa ttaagccgca taatctcatg 20159  
 gatataacat ctcatgggtg atccttaatt aacagcttat atatatatat atatatatat 20219  
 atatatatat atatatatat atatatatat atatatatat tgaccctaac tgtggcaaac 20279  
 atgcattatt atcacacaaa agttactaac cacatatagg agcctatggc taatggctct 20339  
 gagtagaaaa atgggcacag aggatctcca tgatactatt tatggcaact cacgtagcaa 20399  
 aaagccgcag actaacacat ccatggatat ccacaacgca tactgatagt agtctgatat 20459  
 acacactagc tcttcccatg acggccttag cgaaaaccac tttttaaccc aaaaaaaaaa 20519  
 ccagttagga ccggtgaaaa gtcgcacgcg atgatcgatt cacgcgcgcg ccgcagaagc 20579  
 aacttgcaaa agggatcgag cttagctaga tagcgcgagc tcatcagcat ttcgtcgtcg 20639  
 ccgagcgagc tagtggcttt ggcagttagt agtgatggga gttgcataga agttaagaac 20699  
 caggtagaca gagatcgatc gattgatcaa acccgtttgg tttcggataa gtatgggaag 20759  
 aatctgaaac agtgtggagg aaacactgag agagaaaaga caccattaac aataatatcg 20819  
 atggaattcg ttttttttgg tgggtgttgc tagaagccta gaacagcaat tcatgtgatc 20879  
 gatcgatact tcgatcgtgt gcgtgtgtga cgagaaagag atggggcatg tgaaggcaaa 20939  
 gacgaggttg acatttgcac agctagccgt tctctcctga cagaattaag ctagaaattg 20999  
 aagatccgtg actctgagta gtcctaacca attagctata cgcctataca cgatgggcta 21059  
 gctatgcacg cacgcgacgc caaattgaac acggatgaac aaataaaatc gaacaatggg 21119  
 ttggctagcg caatcgatcg atcgatctta ccgttgctgg ccatgaggtt ggagaagaat 21179  
 ccggccggcg agctgctgtg ccgcgccagc aagtccacgc tcccgtcctg cagctgatgc 21239  
 ccatgccctc cgccccctcc gccgcgcgcg ccgcgcgcgc cgtgcggggc cagcgagatg 21299

tccccccac	cgaaccgcag	ccccgcgcgc	tccgcctccc	ttggctgcgg	cgtcgctcac	21359
gacgacgacg	gctccacccc	tcctcctcct	cctcccccca	cgggcagaaa	cctcctcatc	21419
ataccctcc	ttggatcgat	cgatcaactc	cacccccgcg	gaccgagacg	cggcctctcg	21479
tcgatcgatc	tgcagctcgc	gcaggcgacg	gtaggcaggc	ggcgcgtggt	gtgggtggaa	21539
atttcggcgt	gaaaattaac	aaaacgcagg	ggggggccta	tactatagct	agtagaggag	21599
agaagagggg	aaggggaagg	aagggggagg	tgaggtggtg	gaggtggtgg	ggctaggcgc	21659
aagtgggagg	agaggggtgt	gggattttta	aggggaagcg	ggccccgtga	ttggttctcg	21719
gggcgtgtgg	cgccgtgggg	accagcggac	cggccggggc	cgggcaagtg	gatgtctcgc	21779
gcggagtggg	gtgggcttct	gcactgcgca	gcagcagcag	tagcaagccg	taggtggcgt	21839
cgcgcgcccc	gccccggaac	cggcaggcat	ctctctcggc	ttttcgctgc	atctttggtg	21899
ctagattttt	gtgttggata	tatgatgctg	atcgaggaaa	gggaagggaag	aagaaaaaaa	21959
aaaggatttt	tttgggtgtg	cttagatttt	tggatgcttt	ctttcctctg	ctgcggactg	22019
cggggactag	aggatgaact	cgataatcaa	tgggtggtgg	ggcaaagtgt	tatacttcct	22079
cagtctttta	tatttacctt	ttgtgatatg	gaggaaacaa	gctgggttgt	ggtgttgtgc	22139
actcgtagga	ggggaggtac	gtagttaacg	gcaaagatcg	atcatgcaag	ttggttgggt	22199
caatttggtg	gtcgagctga	cctatgttcg	cccatcctct	cgatactttt	ctcatctaga	22259
ctttttctac	gacgctaaca	gactgattat	cacagtcatt	ggatagatcg	acatggtcat	22319
ttgaaattgt	tcgatataac	tggtttaagt	tcaaacaaat	ccaagctaaa	ttttattttg	22379
cggaaaaaat	gtttgaattt	cacttgtttt	caaccgttat	tgctgttagc	gaccttgccg	22439
ttagggagcg	gttttttaac	ctcggcacat	cogtaaacct	tattgcaggg	gagtcattgt	22499
tatgtctaac	agtagtataa	ttttatcaca	atgatttgtc	tctttacgag	ttgtattata	22559
aactcacggt	gttccccgca	aaaaataaaa	aataaactca	cggtatgtgt	aaatggaatt	22619
aggtcaaaat	ttaggaatga	aatgaataat	caattgggtg	tgaatgggtc	aatgcactaa	22679
accatatgtt	ttgctcacta	gatatgacaa	ggaaaaaccg	aaccatcaat	aacactggaa	22739
accatgtttt	tgtgggtgac	cttagttaac	tcatacatca	attataatct	tttctctatc	22799
caattccact	ttggtctatt	ttgtctattt	gaaatcatgt	ttcagctatc	ttctaagtaa	22859
agcaaacttg	aaaacctagt	acatctaaac	ctagctccac	tagtgtgggtc	caaaagcagt	22919
ggagtaatat	cataagagga	agacaacaaa	aaataggata	gagatagtct	tagcttgtgc	22979
cgcagtaatt	cattcgatag	attattagta	attcattcga	taagatatga	taatgatgaa	23039

atgattcggg cgcacgggtg ttgtgaagtc ggagccatga tatgtggcat cgaaagcatt 23099  
 agtcaaacgg acttcgggtt tggtcaggta aagttgtgtt ccttggttct taattcttat 23159  
 caaaattgga gtcgcctgat catgtgtgcg gtgggtgtgat gacgaatgac ggcgagtttt 23219  
 taccaatgta cagtgaactg cgttttgttt taaggctgtt agttttgttg tcgtggttat 23279  
 gctttgctag ctagttaggg tgattcctat tttttgtcag gtcttatgaa agttaaaaaat 23339  
 attttttaaa atatgagttt ttttattgtg aataatgagg aacaaatgaa gttttgggag 23399  
 gatacatggc tagaaaacat ggcttttaaa gataaatatc catctttata ttatatagtt 23459  
 cgaaggaaaa atttatctat tgctaatacc atgggatctg ttccgcttaa tgtttctttt 23519  
 agaagagttt tagttggtca gaatcttgta tattggcatg aattgcgtgc ttctattgta 23579  
 catattcagt tgaatcaatc tactgactat tttagatgga attatcatca aaatggttta 23639  
 ttttctgtaa ggtcaatgta tctagcctta agccttaatt aataatgggtt acattgagag 23699  
 aaataagatt atttagaaac ttaagatgcc gcttaaaatt aagattttta tgtggtactt 23759  
 gcttaaagggt gttatgttaa caaagacaat ttggcaaaac ggaattggaa tggcagctta 23819  
 agatgttggt tatgtatgaa aaatgagact attcaatatc tttttataga ttgtcatttt 23879  
 gcaaaatttg tttggggagc gtttcagtag tcttttggtt tataccttcc tactttcata 23939  
 cattgtatgt ttgatggttg gctttggggg tgaacaagaa aaggagcaaa ctaattcttg 23999  
 tagaagcttg tactatatgt taggctctgt gattgagtag gaatgatatg atttttgaca 24059  
 aatcactatc tatttcattt atgcaggcat tcttcagagc aacatattgg ctccggtttt 24119  
 gggcacaacc gtaaaagtgt gatgaagatg gagagctttt gaaagttata tgtcgtaaagc 24179  
 ttgagacgac ggttatgcaa ctttttgcca actatggatg gagattcaca aatagactta 24239  
 aataattgtg tgctccttat attgggtctgg tcattttttt tatgttttagg tgtgtgttta 24299  
 attctatttg aactacactc tttgttaagt gctagattgt aataattggc tgtagctctg 24359  
 ttgagcaaag gccgagatgt tatctattcc attattaaaa aaaagctagt tagtgtattt 24419  
 attgttgtat ggcgggtttta gcccgattgt tctaaatcaa ctgaatatta atttgctctt 24479  
 ttttagagaa acaccagag gtcttcgggc tgggttagat gaccttggtc cttatccctt 24539  
 ctaattattt gatattaggt acttcactaa tattcgtatc ttttttaaat taatttgctc 24599  
 tctttttaaa ctattcatct tttctttaat atagcactaa attaaccgtg atctttcaaa 24659  
 agaaaaggcg aaagggtgtga atatgcatga aagatcgagt ggacaccccc caaaaaaaaa 24719

aaaccctagt tgttgtcacg tgactctcaa agtccatttg aggacttact aactgtttga 24779  
 aattaatgga taaggctcca gctaagtagg cgggaaaaga tcaaacgtgt tcagtggatt 24839  
 tataccaaat gtggtccgtg cgacatgttg gtccataaaa gggcatatga aagtttcctt 24899  
 tcagctaatt aaagccagtg tcgaatactt atacagtata gttttcgaaa taagttttac 24959  
 ttctacaatg taatccattc acggatgaaa aagctgtgcg cccaacagct atagctatac 25019  
 aactatatct atttgtaaat taagaggttc atatcttggg cacacaaagg ttctgtttga 25079  
 atcttctgaa gataaatatg aagatcaa atgtttacgta aaacgaggtg gtaataactt 25139  
 ttgattaatt agattttaat tattacaaac ttaaaaaaaaa gattaatctg atattttata 25199  
 acaactttca tatagaaaat tttcacacga aacgcaccgt ttaacagttt gaaaagcgtg 25259  
 ccacgaaaat ctagaactta atctgccctt tgttggttgc tcgaacagga ccaaacttca 25319  
 tgtccatact ccgtactgta cataccaact atactaaata tcgctaaaac gttttaaaaa 25379  
 tattatacat atactttcaa tactattata cgtatgcgta aagttttatc ctcaaattca 25439  
 ttatatattca tacttaaaaa aaattcta atgtttatga atataagtct tagatttttt 25499  
 tctccatata tatatatgat aaatttaaag atgggacttc acgcgtatat ataaatacta 25559  
 tttaaagtac atgtacattt ttctaaaaaa ataatatattg ttagtttgta tacattgtgt 25619  
 gtatacgtga aggtcacgt agacattttg cactctcaat tatttatact agactaataa 25679  
 ccacctaaat attgttctta gcggttttga cttgagctta cctacgagat gccaacgtgt 25739  
 cagtccagtc agcaaaaaag ttttaaaaaa actccgtggg ccacttgtc atacttctcc 25799  
 ctcaatctaa cgctccccg gtcacctac tctctcttcc tcgtgcgcac gctcgtgcgg 25859  
 ccaacggcgg agtggtgcgg tgttggtggt ggtggtgcgt ccgcgtcggg cacgacggcg 25919  
 gtgttgccgg agagatggag ctacgcaaga ggcgggcgtc gacaagtatg gacccagga 25979  
 ggtggaggct acgcgcgagc gcgtcggccg cgcccttgc cttcaacggc gacgacagca 26039  
 ttggccgctc gttctcggcc tcgcctctct gaccgcaggt cggcctcagc ctagcagtag 26099  
 tcggccacgc acgttggcct catcttcgtc accgtgttct tggggctggc atgcaggcgg 26159  
 gagaaggagg gatagcggca tggactgcgt gtgcgcctgt gcagtgcctt ggggtggatac 26219  
 tatgtcaagt tggagctctg caagtcggcg ctctgcgacg gcgacggcaa caggacaca 26279  
 tcgtcgttgt caccgtgctg cgcggcacga gaaagatggt ggcactgacg gtgaggcttg 26339  
 cgatgacaaa tgtgatggtg acaagggccc aacgtcgagg tcgttgccct ggaatagaat 26399  
 ccgatcggca gctctagcgg tggcaacggc tcagtgcctg cgctagagta cgagcgcggt 26459

cggtggccac agggcggtac cgcggcgcgc gggaggtgct ccaggcggcg cagtgtcccc 26519  
cccgtgagaa cgagcggttt agcatcatgg ccatcgtcgc gcaccgtcac cgccctgggc 26579  
ttctctagcc gaccacgctc cacctcaagc aaccgggga gtagtctttg ctgccaccgc 26639  
ggccttctcc tcaactgctcg gttcaaggat gagagagaga tcgaggtgga aggaggtaga 26699  
agagatgagg tgagaatata tggatcactg acaagtgggc ctgttatttt ttgccgcgtt 26759  
agaaatgcca agtcagctaa cctagcctaa aaccgtccaa aatagtgcc cggtattcgt 26819  
ctggttttta gagttttgag gtattgaata caatatatgt tattatagtt tagagggtaa 26879  
attgtactac cgtaccataa tagttcgggg gtaaattgta cttcctctgt actcataatg 26939  
gaagtcgttt aggacaatat ttaagtcaaa cattgggaat ataaatcatg aataactctc 26999  
aagttgttga gtttgaaaat gtaaaaatta tatgaataga tttttcttga aaaatatttt 27059  
cataaaagta tacatatatc actttttaat atatatatttt atagaaacaa gaagtcaaaa 27119  
ttatgttttg gagaccgtgt cgctgtccaa aacgagtacg gagggaatac tttttactcg 27179  
tagtttacia tatcgatctg ttaactgttt ataagagtat ttggatccat gcagtattgt 27239  
agtagtagta gcagtacatt tgagaatatt agagtacgaa ttaggtggtg tttggataca 27299  
gagacttaac ttttagtcttt gtatttagac actaatttag aatattaaat atagactact 27359  
tacaaaacta attatataaa tgaaagctaa tttgcgagac aaatTTTTta agcctaatta 27419  
atctataatt agagaatgtt tattgtagca tcatataggc taattatgga ttaattaggc 27479  
tcaaaagatt tgtctcgtga attagtctaa gattatgaat gagttttatt aatagtctac 27539  
gtttaatatt tataattaat tttcaaacat ctgatgtaat agggacttaa aagactttta 27599  
actaccattt aaacagggtc actccaatgg taggtgaaat tcaacagctg ggaaatgcac 27659  
tagtgcgttg tgtcagtaaa tttcgtacta gtaccacgag acagctagac agacacgtca 27719  
ggtcacgacg cagcactgca gcagggtgt agcctgtacg ggaggcgtag gcgcaacatc 27779  
tcgaaaattt tgttccgtag ctaaagcccc cccaaagcca gccgcggttt tcatggattg 27839  
cacaggcggt cctctccgcc ggattccgga aagaaaaaag aaaaaacaag atgtccgttc 27899  
cctgggtggt gcatccgttt tctgacaggt gcatgcacct ctcgctcgct accgcggtag 27959  
cgccacacg aaccacgttg gctttcggcc aacttgcccg attctttaat cccctcacga 28019  
cgtacgtcgc tgtccaataa aaagttttta caccaactat agtaaccagc ttaattttta 28079  
taaaacaaa gaaaattctt aattacttag agcatctcca acagggtcct caaacaaggt 28139

ccctaaataa gttttgagag ttgatgcaaa aaaatatagg tccagcagat tccctactag	28199
agcccccaat ctagggaggg ccctagatca ctctccaag cccccagtcc gggggggtca	28259
accccacagc ccccatcctc tttttttttg gcgggggaaa tttctgagcg cgcgccatcg	28319
tcatectccc tcccgcgtca tcgccatcct cccacaacct cccagcgagc aagccgccag	28379
gtcgttttgc tccggtcggc gaccacccga catccctcgg gaagaaaactt cggcgagatc	28439
ccatggtgtg ccgcctcaag ccatcacaat catcggtttt actccgtcac tcaactgtcg	28499
cgacctcctt gcatcttcgg ccaacgcagc tcgtcaccga ttgatcatgc ccaggtgagt	28559
ctcgacgtcc tccccgaact agctgccac ccgctgtcca caaggcctag cgccgaccat	28619
cgacagctat ctatcaccgg ccgcacacat gctgcagtaa aaaaattcag agatgatggt	28679
gattaaaaca aatcacaata gtaaagttca gtttcgtatg tctgagtctc cttgtttgat	28739
tttgatcttt atgggcttac taggcgtcta ggccatcta aatcattcgc acagcaaac	28799
gtacattgtc atcattcatg ttttatatgc agtgtcttgt tctatgtcag agagctaate	28859
ttgcagagca tatataatat ttaagaaata aatttgtgtt tgcaactgagt ccttagtact	28919
gcgcaaccaa tatatatgct aaataaatac atattgcaaa cagtataacc tgatgtacat	28979
tgcaatcact tgttgatgtt tctgagatag attggaaagg ttgtcaattt atatatttat	29039
tgcaactgact agattgcaat gacaagtgga ggtgattcct ttgtgcgcat gatgtccgag	29099
gacactgatg tcgaagtgtt aatgccaaat gaagaccttc gtacttcaac aaatggtgca	29159
aaaggaagtg ccaaaagatc aagcaactat actcataagg aggacattca attgtgcatt	29219
tcattggcaga gcattagctc agatcctatt attggcaatg agcaaccagg gaaggcatat	29279
tggcagagga tcgcagagca ctaccatgct aaccgtgatt ttgagtctga taggaatgca	29339
aactctcttg agcaccattg gggtaacatt cagaaggaag taagcaagtt tcaaggttgc	29399
tacaatcaaa ttgagcgctg tcatccaagt ggcataccac atcaagagct tgtaagttaa	29459
attgtttatt tattattatt aataacaatc ttgtatgtat gtgaattaaa acttaaatta	29519
tgttgcaggt tcttgaagct gaggcattat actcgtccac tgcaccaaag aatagggcat	29579
ttcagtttaa tcattgttgg ctcaagttga ggaattctcc aaagtttcaa aactagaat	29639
cccacaagag gccacggtct aggaagtctt cgacccaat tgagagagct ggtgaagaag	29699
atgaaggaga tgatgctagc aagagtacag ctctgattt atctcagccg agtgctaaaa	29759
agagaccaat aggtaggaag caagcaaagg aaaagttgaa gaatggagga caagatggac	29819
catacaaaga ggcgatgaaa gatttgcttg acgctaaaga gaaagaagcg aaattgaaa	29879



aagagagatg	gaaggaaact	aaggagattc	aagagcgcaa	gctcttattt	gctgagcgtg	29939
agttagtgtg	ggatcaagaa	cagaagatta	tgttttgtga	tgtttccacc	ttggaaccgg	29999
atgtgagaac	gtatgtgttg	gctatgaggg	cacagattgc	agcttcaaag	gtggctgccc	30059
tcaatggtgg	atttgatggt	agtagtggct	ttggaggtga	gtttggtggc	ggtaatggag	30119
aagtttgagc	acttcgatgg	aataagttgg	attctattgg	atgatccatg	tgtcctttac	30179
tagtaggata	tgccattatc	acgattggtc	tttggagtcc	ttttttgtta	attattttcca	30239
caataatttt	agtgtcactt	gctagtagga	catatattac	tttcagattt	gttattttata	30299
atcgaatcat	tcattggtgt	aggatgtatt	attttttaa	tatataatgc	atcattgggt	30359
tcacatagt	gtatgtgttg	gagcaatttt	cattttcatt	ggtgaattac	gaatcttggg	30419
tgcatcttgt	tgctgtatat	ggcactgtac	ccataccata	tttacctgtt	taaaaatttt	30479
aattttgtat	tcgaattgta	gtgtttgaaa	ttgtgaattt	aagtatgggt	aaattatgtg	30539
agttagaaat	aattgtgttc	gaatttttgt	ggtgttaa	atactgtata	tggattgtat	30599
tttaaaatac	aagataaaca	tgagtaggga	ctaagaaata	ggggctactg	ctggagttgg	30659
aggcattttt	tagtccttga	gaaatggggg	cagccctcat	ttaactttta	gacgcttcaa	30719
aataaggtct	attgctggag	atgctcttag	gtcccatcgc	tttccttcaa	tcagcattag	30779
ccgctaccaa	aatttgaaat	tttaaagttt	ttcatcgaag	tttattttcc	agcattggta	30839
tttaagtcgc	taaaaacaca	tatatgaaag	tcttatctgt	aaattattat	tattttgcta	30899
atacgccgaa	tggcgtatta	tatgtatttg	gccaaaggat	gggggcctta	aaccttagcc	30959
ttagtcgtgc	cctacaaaag	acacacgcct	cgtcagggca	agggtactcg	agcgtggagg	31019
catggttcgc	aagccatggt	cggcgaggcc	atgctctagc	aatgcggtgc	agtccacctc	31079
ctctccgagc	gcggagctcc	aacgggtgat	ggccaatgaa	agaaggagac	cgacttgccg	31139
ttggtttag	catgtaaatt	tcttgcaact	tcttaataaa	tttcggctag	tgttcgctag	31199
ctcgacaaaa	aaaaagagag	gctaattgat	gggttaggaa	gtgaaaacaa	gcgcagtggg	31259
ggagaagaag	atcgagaggc	ctatttgat	gatgctttgt	cgatgtagat	ttagtcccat	31319
gctcatctca	tccctcagcc	acaacaatcc	catcattgta	gagctcatca	gcttgctcta	31379
ccatctctcc	ttgtttatgg	gccactccca	acctgctacc	catcgctga	tctatgaatc	31439
tagctgtcaa	tgacctcatt	ggccctagtc	ttgagatcac	ccagtggatc	ctttggcaaa	31499
gtggatccgc	ctttgttttg	ctttggagaa	agaaaacgat	gacttagcta	aagatctcgt	31559

cgggtcaaaaa gagagatgcc tttgatatat gctgaaaaat agaggagagg cagtgtcagc 31619  
 tggagagctc tttatccaca cccgtgggga tcgagcttat tggcgtagag ggagagacat 31679  
 tgagggagag agagtgaag gggatttttt tgtaatttct agatttggtg gtgttttagtg 31739  
 caatactttg aacttatttg taaattaagt aaaacatgat tgtaatagaa aatatcataa 31799  
 actgacatag aaaaacaaag ataacaattg aagccactag cgctatggag aaaatgtgtg 31859  
 acctcggctc acatataacg gctatgtgtt attaccatgt cacttctaaa actaccatat 31919  
 aaccatatac gtttttctcc tacttatcaa aaatataatt aacaaatttt ttaccgggtt 31979  
 tagtttacia gaaaaaagtt tgactgcatt gttgataccc taccatcctt gtacgaaggc 32039  
 aggcgctaca caacaccgct gccgctgccg tcgcgcgcgt aagctaaggc tgtcacgccg 32099  
 gcgaccggcc acggccgacg tggaaagcga cctaactctg aaagtgtaaa cccaccctat 32159  
 agaaaaaccg ggttggtggg acgagaatca ccgaatcagc gtcgacgacg acggccgacg 32219  
 actccagcag cgggggtcac gagactcgga gccgagagag agaaagagga ccacgcgcgc 32279  
 attcactcaa ctgcataaaa aaacccccgc gggcggtgtg cgcagtcacg tctacgctcg 32339  
 cgggatcgct cgatgaaatc aaccaaatac ttaaacaac cgaaccaacc aaccaaccgt 32399  
 cgcgcggtgt cgcgcgaggc gctcgattag cggagacgca aacccatgta acaccgtgcg 32459  
 gaaaaactta aagaaatccg cgtcgctcgc gccgtcgcgc gcgcgggggg cgcgtagtac 32519  
 ctccacacac gattctgcac ttgtactacc acgcgaacct gatgcgggtt accggtcac 32579  
 gattggctgc gaggcttgct gttactggtg gtggtagact ggtagtacgt tgcttgact 32639  
 acctcactca tgtctggaga ttactacact tcgatctttt cctctgtttt gtttaattgag 32699  
 atttgagggt gttactgttc gctgtgtggt taagtatat ggtgtataac tacaagttgg 32759  
 tactctcaaa gggaaaaaaa ggtactgcaa attggctaatt ctatgattct attctgcaca 32819  
 tgcatataga taagcactat aataaggaac tgaggatcgt gaaaagtggc attaatata 32879  
 acaggaccat gtacgactat accactggca gggatttcac ggaatcaact ataggagtag 32939  
 gttagttggc acttggcaag gttgattgat tcactaacgt ggggaaaaga acacacgaga 32999  
 tcaaaggctg tcgtgggctt aaaataaaag ggccatctg ggatcagctc ttttaagccc 33059  
 acatcactag ccaggaggct aggagtcag tattgcctcg tactgggccg tcctctgaaa 33119  
 tttggaggcc ctgtctaaaa ttctaataaa gccttaaaact taagtgacaa aataaaaaga 33179  
 ggtagactat ataacagcat accattacaa cggaatagct gtcgttagca cgatactcta 33239  
 tatgcatcag atatggtacc aggtactata ccgacgttag catgatccga taggtatagg 33299

```

atctggtgta cctagatatt atgctaacat aatcatgaca tcagctattc cattggaatg 33359
atataccggt ggtatcttcg gtaaattgtg agcatgctag gaatttaagt aaagggcctt 33419
agggttaaaa tcacacgttc ttagtcaactg cactatcaag tgcatttcaa ccctaagtc 33479
cttttatgat ctatatctgc cctcctagcc tattttggac gaggctccct cgtcctagaa 33539
gtaaatcatc gtatccataa tccaaccgat tagtagagaa aaaacatact tttcgaacgc 33599
aacagttctt gtcattctgt gctctcaaat gttcattttc cccttactta aaggacatgg 33659
aaaacagaac agaccc 33675

```

```

<210>      3
<211>      1119
<212>      DNA
<213>      Escherichia coli

```

```

<220>
<221>      CDS
<222>      (1)..(1119)
<400>      3

```

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400

```

atg cat aac cag gct cca att caa cgt aga aaa tca aca cgt att tac 48
Met His Asn Gln Ala Pro Ile Gln Arg Arg Lys Ser Thr Arg Ile Tyr
1 5 10 15

gtt ggg aat gtg ccg att ggc gat ggt gct ccc atc gcc gta cag tcc 96
Val Gly Asn Val Pro Ile Gly Asp Gly Ala Pro Ile Ala Val Gln Ser
20 25 30

atg acc aat acg cgt acg aca gac gtc gaa gca acg gtc aat caa atc 144
Met Thr Asn Thr Arg Thr Thr Asp Val Glu Ala Thr Val Asn Gln Ile
35 40 45

aag gcg ctg gaa cgc gtt ggc gct gat atc gtc cgt gta tcc gta ccg 192
Lys Ala Leu Glu Arg Val Gly Ala Asp Ile Val Arg Val Ser Val Pro
50 55 60

acg atg gac gcg gca gaa gcg ttc aaa ctc atc aaa cag cag gtt aac 240
Thr Met Asp Ala Ala Glu Ala Phe Lys Leu Ile Lys Gln Gln Val Asn
65 70 75 80

gtg ccg ctg gtg gct gac atc cac ttc gac tat cgc att gcg ctg aaa 288
Val Pro Leu Val Ala Asp Ile His Phe Asp Tyr Arg Ile Ala Leu Lys
85 90 95

gta gcg gaa tac ggc gtc gat tgt ctg cgt att aac cct ggc aat atc 336
Val Ala Glu Tyr Gly Val Asp Cys Leu Arg Ile Asn Pro Gly Asn Ile
100 105 110

```

ggt aat gaa gag cgt att cgc atg gtg gtt gac tgt gcg cgc gat aaa Gly Asn Glu Glu Arg Ile Arg Met Val Val Asp Cys Ala Arg Asp Lys 115 120 125	384
aac att ccg atc cgt att ggc gtt aac gcc gga tcg ctg gaa aaa gat Asn Ile Pro Ile Arg Ile Gly Val Asn Ala Gly Ser Leu Glu Lys Asp 130 135 140	432
ctg caa gaa aag tat ggc gaa ccg acg ccg cag gcg ttg ctg gaa tct Leu Gln Glu Lys Tyr Gly Glu Pro Thr Pro Gln Ala Leu Leu Glu Ser 145 150 155 160	480
gcc atg cgt cat gtt gat cat ctc gat cgc ctg aac ttc gat cag ttc Ala Met Arg His Val Asp His Leu Asp Arg Leu Asn Phe Asp Gln Phe 165 170 175	528
aaa gtc agc gtg aaa gcg tct gac gtc ttc ctc gct gtt gag tct tat Lys Val Ser Val Lys Ala Ser Asp Val Phe Leu Ala Val Glu Ser Tyr 180 185 190	576
cggtttgctggcaaaaagatcagatcagccgttgcatctggggatcacc Arg Leu Leu Ala Lys Gln Ile Asp Gln Pro Leu His Leu Gly Ile Thr 195 200 205	624
gaa gcc ggt ggt gcg cgc agc ggg gca gta aaa tcc gcc att ggt tta Glu Ala Gly Gly Ala Arg Ser Gly Ala Val Lys Ser Ala Ile Gly Leu 210 215 220	672
gggtctgctgctgtctgaaggcgtcggcgacacgctgcgcgtatcgctg Gly Leu Leu Leu Ser Glu Gly Ile Gly Asp Thr Leu Arg Val Ser Leu 225 230 235 240	720
gcggccgatccggtcgaagagatcaaa gtc ggt ttc gat att ttg aaa Ala Ala Asp Pro Val Glu Glu Ile Lys Val Gly Phe Asp Ile Leu Lys 245 250 255	768
tcgctgcgtatcgttcgcgagggatcaaac ttc atc gcc tgc ccg acc Ser Leu Arg Ile Arg Ser Arg Gly Ile Asn Phe Ile Ala Cys Pro Thr 260 265 270	816
tgttcgctgcaggaatttgatgttatcggtagcggttaaacgcgctggag Cys Ser Arg Gln Glu Phe Asp Val Ile Gly Thr Val Asn Ala Leu Glu 275 280 285	864
caacgcctggaa gat atc atc act ccg atg gac gtt tcg att atc ggc Gln Arg Leu Glu Asp Ile Ile Thr Pro Met Asp Val Ser Ile Ile Gly 290 295 300	912
tgcgtggtgaatggc cca ggt gag gcg ctg gtt tct aca ctc ggc gtc Cys Val Val Asn Gly Pro Gly Glu Ala Leu Val Ser Thr Leu Gly Val 305 310 315 320	960
accggcggcaac aagaaaagcggcctctattgaagatggcgtgcgc aaa Thr Gly Gly Asn Lys Lys Ser Gly Leu Tyr Glu Asp Gly Val Arg Lys 325 330 335	1008

gac cgt ctg gac aac aac gat atg atc gac cag ctg gaa gca cgc att 1056  
 Asp Arg Leu Asp Asn Asn Asp Met Ile Asp Gln Leu Glu Ala Arg Ile  
                   340                  345                  350

cgt gcg aaa gcc agt cag ctg gac gaa gcg cgt cga att gac gtt cag 1104  
 Arg Ala Lys Ala Ser Gln Leu Asp Glu Ala Arg Arg Ile Asp Val Gln  
                   355                  360                  365

cag gtt gaa aaa taa 1119  
 Gln Val Glu Lys  
                   370

<210> 4  
 <211> 686  
 <212> PRT  
 <213> Oryza sativa

<400> 4

Met Ala Thr Gly Val Ala Pro Ala Pro Leu Pro His Val Arg Val Arg  
   1                  5                  10                  15

Asp Gly Gly Ile Gly Phe Thr Arg Ser Val Asp Phe Ala Lys Ile Leu  
                   20                  25                  30

Ser Val Pro Ala Thr Leu Arg Val Gly Ser Ser Arg Gly Arg Val Leu  
                   35                  40                  45

Val Ala Lys Ser Ser Ser Thr Gly Ser Asp Thr Met Glu Leu Glu Pro  
                   50                  55                  60

Ser Ser Glu Gly Ser Pro Leu Leu Gly Ile Thr Arg Arg Leu Leu Phe  
   65                  70                  75                  80

Thr Leu His Met Val Gly Asn Val Pro Leu Gly Ser Asp His Pro Ile  
                   85                  90                  95

Arg Ile Gln Thr Met Thr Thr Ser Asp Thr Lys Asp Val Ala Lys Thr  
                   100                  105                  110

Val Glu Glu Val Met Arg Ile Ala Asp Lys Gly Ala Asp Phe Val Arg  
                   115                  120                  125

Ile Thr Val Gln Gly Arg Lys Glu Ala Asp Ala Cys Phe Glu Ile Lys  
                   130                  135                  140

Asn Thr Leu Val Gln Lys Asn Tyr Asn Ile Pro Leu Val Ala Asp Ile  
   145                  150                  155                  160

His Phe Ala Pro Thr Val Ala Leu Arg Val Ala Glu Cys Phe Asp Lys  
                   165                  170                  175

Ile Arg Val Asn Pro Gly Asn Phe Ala Asp Arg Arg Ala Gln Phe Glu  
                   180                  185                  190

Gln Leu Glu Tyr Thr Glu Asp Asp Tyr Gln Lys Glu Leu Glu His Ile  
 195 200 205

Glu Lys Val Pro Asn Ile Ser Leu Phe Ser Val Asn Leu Val Phe Ser  
 210 215 220

Pro Leu Val Glu Lys Cys Lys Gln Tyr Gly Arg Ala Met Arg Ile Gly  
 225 230 235 240

Thr Asn His Gly Ser Leu Ser Asp Arg Ile Met Ser Tyr Tyr Gly Asp  
 245 250 255

Ser Pro Arg Gly Met Val Glu Ser Ala Leu Glu Phe Ala Arg Ile Cys  
 260 265 270

Arg Lys Leu Asp Phe His Asn Phe Val Phe Ser Met Lys Ala Ser Asn  
 275 280 285

Pro Val Ile Met Val Gln Ala Tyr Arg Leu Leu Val Ala Glu Met Tyr  
 290 295 300

Asn Leu Gly Trp Asp Tyr Pro Leu His Leu Gly Val Thr Glu Ala Gly  
 305 310 315 320

Glu Gly Glu Asp Gly Arg Met Lys Ser Ala Ile Gly Ile Gly Thr Leu  
 325 330 335

Leu Met Asp Gly Leu Gly Asp Thr Ile Arg Val Ser Leu Thr Glu Pro  
 340 345 350

Pro Glu Glu Glu Ile Asp Pro Cys Arg Arg Leu Ala Asn Leu Gly Thr  
 355 360 365

His Ala Ala Asp Leu Gln Ile Gly Val Ala Pro Phe Glu Glu Lys His  
 370 375 380

Arg Arg Tyr Phe Asp Phe Gln Arg Arg Ser Gly Gln Leu Pro Leu Gln  
 385 390 395 400

Lys Glu Ala Pro Glu Leu Leu Tyr Arg Ser Leu Ala Ala Lys Leu Val  
 405 410 415

Val Gly Met Pro Phe Lys Asp Leu Ala Thr Val Asp Ser Ile Leu Leu  
 420 425 430

Lys Glu Leu Pro Pro Val Glu Asp Ala Gln Ala Arg Leu Ala Leu Lys  
 435 440 445

Arg Leu Val Asp Ile Ser Met Gly Val Leu Thr Pro Leu Ser Glu Gln  
 450 455 460

Leu Thr Lys Pro Leu Pro His Ala Ile Ala Leu Val Asn Val Asp Glu  
 465 470 475 480

Leu Ser Ser Gly Ala His Lys Leu Leu Pro Glu Gly Thr Arg Leu Ala  
 485 490 495

Val Thr Leu Arg Gly Asp Glu Ser Tyr Glu Gln Leu Asp Leu Leu Lys  
500 505 510

Gly Val Asp Asp Ile Thr Met Leu Leu His Ser Val Pro Tyr Gly Glu  
515 520 525

Glu Lys Thr Gly Arg Val His Ala Ala Arg Arg Leu Phe Glu Tyr Leu  
530 535 540

Glu Thr Asn Gly Leu Asn Phe Pro Val Ile His His Ile Glu Phe Pro  
545 550 555 560

Lys Ser Val Asn Arg Asp Asp Leu Val Ile Gly Ala Gly Ala Asn Val  
565 570 575

Gly Ala Leu Leu Val Asp Gly Leu Gly Asp Gly Val Leu Leu Glu Ala  
580 585 590

Ala Asp Gln Glu Phe Glu Phe Leu Arg Asp Thr Ser Phe Asn Leu Leu  
595 600 605

Gln Gly Cys Arg Met Arg Asn Thr Lys Thr Ile Ala Ile Met Gly Cys  
610 615 620

Ile Val Asn Gly Pro Gly Glu Met Ala Asp Ala Asp Phe Gly Tyr Val  
625 630 635 640

Gly Gly Ala Pro Gly Lys Ile Asp Leu Tyr Val Gly Lys Thr Val Val  
645 650 655

Gln Arg Gly Ile Ala Met Glu Gly Ala Thr Asp Ala Leu Ile Gln Leu  
660 665 670

Ile Lys Asp His Gly Arg Trp Val Asp Pro Pro Val Glu Glu  
675 680 685

<210> 5

<211> 594

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> unsure

<222> (1..594)

<223> unsure at all n locations

<400> 5

aaaatcgtca atccctctca aactcttctc accactaatt tcttcctctg gaacattctc 60

ttctctatta ttttgattcc cttggcctca acactgggtt ctcaattgca tgatcttggc 120

tcgtcttcag ttactttgat tcactgagaa aaatggcgac tggagtattg ccagctccgg 180

tttctgggat caagataccg gattcgaaag tcgggtttgg taaaagcatg aatcttgtga 240

gaatttgtna tgtaggagt ctaagatctg ctaggagaag agtttcggtt atccggaatt 300  
 caaaccaagg ctctgattta gctgagcttc aaccctgcat ccgaaggaaa gccctcttc 360  
 ttagtgccaa ggcaggaaat attgtgaatc attgcataa gcggttagga ggaagnctcg 420  
 gacctgtaat gggtgaaatg tcgncccttn gaagnnaca ccggtanggg tcaaacggtg 480  
 ccttcttngg gtacaaaang tnttccttgg ancctnttng tggggggttt gggattgctg 540  
 aaaaaggggc tgnttttnaa ggnnacctnn caagnagna agggnggggc tttt 594

<210> 6  
 <211> 615  
 <212> DNA  
 <213> Glycine max

<220>  
 <221> unsure  
 <222> (1..615)  
 <223> unsure at all n locations

<400> 6

accagaagtg atgagcctta tgaagaactg gacattctta aggggtgttga tgctactatg 60  
 cttttccatg accttcctta tacagaagac agaattagca gagtgcacgc aaccagacgg 120  
 ttatttgagt acctatctga caattctcta aacttccttg ttattcacca tattcagttc 180  
 ccaaattggga ttcacaggga tgacttggtg attggtgctg gttctgatgc tggagccctt 240  
 ctgggtgatg ggcttgagga tggactactt ttggaagccc cggacaagga ttttgaattt 300  
 attagaaaca cttctttcaa tttgttgcaa ggctgcagaa tgagaaatac aaagacagag 360  
 tatgtctcat gtccatcctg tggcagaaca ttgtttgatc ttcaagaagt aagtgcacaa 420  
 attcgggaga agacatcaca cctncctggt gtttcgattg caatcatggg atgcattgtt 480  
 aatggaccag gggagatggc tgatgcagac tttgggtatg tgggaagcac tccccggaag 540  
 attgacctct atgttgggaa gactggtgtg aagcgtggga attcaatgga gcatgccaac 600  
 catggcttga tccga 615

<210> 7  
 <211> 589  
 <212> DNA  
 <213> Lycopersicon esculentum

<400> 7

tggcgatgaa tcacatgatg agttggaaat cctgaagagc tctgatgtta caatgattct 60



tcataatctg ccatatacag aggaaaaaat tggcaggggt caagcagcca ggaggctttt 120  
 tgagtatctt tccgagaatt ccttgaactt tccagtgatt catcacatac aatttcccag 180  
 caacacccac agagatgact tagtgattgg tgccgggaca aatgcggggag ccctcttggt 240  
 agatgggctt ggtgatggac ttctcttgga agctccagac aaggattttg attttctcag 300  
 aaatacatct ttcaatttgc ttcaagggtg cagaatgcgg aacacaaaaa cggaatatgt 360  
 atcatgcca tcctgtggca gaactttatt cgatcttcaa gagataagcg ctcaaattag 420  
 agagaagacg tcacacttgc ctggtgtttc aattgccatc atgggttgca ttgtgaatgg 480  
 acctggggag atggctgatg ctgactttgg atatgttggt ggtgctcctg gaaagattga 540  
 cctttacgtc ggcaagacag tggtgaaacg ccctattgaa atggagcat 589

<210> 8  
 <211> 617  
 <212> DNA  
 <213> Mesembryanthemum crystallinum  
 <400> 8

gaaaagcata gacattatct tgactttcaa cgtagaactg gtcaattacc gattcagaaa 60  
 gaggggtgaag atgtggacta tagaggtgtc ctacaccgtg atgggttctgt cctcatgact 120  
 gtttccttgg acatgttgaa gacacctgaa ctccctttaca agtcattagc agcaaagctt 180  
 gttgttgga tgccatttaa ggatctggct actgtagact ctatctttct gagagagctt 240  
 tcaccagtag atgactctga tgctcggcta gctctgaaga ggttaataga tataagtatg 300  
 ggtgtcatag ctcccttttc tgagcaactg acaaagccct tgccaaatgc aattgtattg 360  
 gtgaacctta aagagttgtc aaccgggtgca tacaagcttt taccagtagg aaccgcttg 420  
 gcagtatctg tgcgaggtga tgaaccatat tgagacattg gagatcctta aagatattga 480  
 tgcttcaatg gctttttatg aactgtcttt taccgagagg atattcacac agtgcattgct 540  
 ggaccaaagc ttttgaggtc ctatcagata agcttggaacc tcccgtaatt aacatatacct 600  
 atcccttcgg attaagg 617

<210> 9  
 <211> 416  
 <212> DNA  
 <213> Oryza sativa

<220>  
 <221> unsure  
 <222> (1..416)  
 <223> unsure at all n locations

<400> 9

```

ggattcggca cgagtctaatt tgatggtctt ggtgatggtg tacttcttga aagctgctga 60
ccaagaaatt tgagtttttg agggacacat cctccaactt gttacagggc tgcaggatgc 120
gcaacacaaa aacggaatat ttccctggtc ctccctggtg gcggacacnc tttnaccncc 180
aaaaattcan tgctcaaatt aaanaaaaaa ccnctcatct gccaggcntc tctattgcta 240
tcatgggtng cattgtcaat gggccagggg aaatggccaa tctaattnc ggataactng 300
gaggtgccct ggagaaaatc nacctntatn ttggttnttt tttttnnaac ggggcatngc 360
aanagaaggg ggcccnacc ccnanatncn ttcnccgggn ccngggccgn ggggtt 416

```

<210> 10  
 <211> 621  
 <212> DNA  
 <213> Zea mays

<400> 10

```

gaattcggca ccagaagcca ctcccacatg caattgtact tgtcaacctc gacgaattgt 60
caagtgggtg acacaaaactt ttgccagaag gcactagact agctgtcact cttcgtggtg 120
atgaatcata cgagcagcta gatattctta aggatggtga tgatataaca atgttggttac 180
ataatgttcc atatggtgag gagaagacag gcagggtgca tgctgctagg aggttatttg 240
agtacttaca ggccaatggc ttgaacttcc ctgtaattca tcacataaat ttccctgaaa 300
ccattgacag agatggtctt gtcattggtg ctggggccaa cgttggtgct ctcttagtcg 360
atggtcttgg tgatggtgta ttccctgaag ctgctgacca ggaatttgag tttctgaggg 420
acacatcttt caacttgctc caaggttgca ggatgcgcaa cacaaaaact gaatatgtgt 480
cttgtccttc ctgcggccga aactctttg accttcagga aatcagcgct gagattagag 540
aaaagacctc tcacttgcca ggtgtctcga tcgctatcat gggctgtatt gcaatggacc 600
aggagagatg gctgatgccg a 621

```

<210> 11  
 <211> 601  
 <212> DNA  
 <213> Pinus taeda

<220>  
 <221> unsure  
 <222> (1..601)  
 <223> unsure at all n locations

<400> 11

```
aatgcaagaa gtacggaagg gcaatgcgaa ttggcacaaa ccatggaagt ctttccgata 60
gtactatgag ttattatggt gattctccca ggggtatggt ggaatcagca tttgaatttg 120
cacgcatttg ccggaagttg ggttttcata attttgtgtt ttcaatgaaa gcgagcgata 180
ctgtagtcat ggttcaggca taccgtttac ttgttgcgga gatgtatgtg caaggatggg 240
attatccatt gcatttagga gttactgaag ctggtgaagg tgaagatgga cgcatagaat 300
ctgcaattgg cattggaaca cttttgcagg atgggttggg tgatactatt cgagtttccc 360
ttacagaacc tccagaagag gagatcaatc cctgtagaag acttgcaaat cttgggatgc 420
aagctgcaaa gctanggaaa ggagtggctc cttttgagga gaacatcgtc attactttac 480
tttccaacgc angactggcn agctccagta cagaaggagg gtgatgaggt ggatacagag 540
gagtccgcat cgtgatggtc tgttctaata tcagtgtcct tgacagntga agacacanaa 600
a 601
```

<210> 12  
 <211> 443  
 <212> DNA  
 <213> Physcomitrella patens

<400> 12

```
gcacgtatct gccgcaaaca tgactatatt aatttcttgt tttctatgaa agcaagcaat 60
ccggtcgtaa tggttcaagc atatcggctt ttagtatctg agatgtatgt gaacaactgg 120
gactacccat tacatcttgg tgttactgag gctggagagg gagaggatgg tcgcatgaag 180
tcagctatcg gcattggtgc tttacttcag gatggtctcg gtgacacat acgtgtttca 240
ttgacggaag ctctgaaga agaaattgat ccttgcacaa agcttgcaaa ccttggcatg 300
aagatttctg cagaacagaa gggggtggct gaattcgaag agaagcaccg gcgatacttt 360
gacttccaac gaaggaccgg ccaacttcca ctgcagaggg agggagagtt ggtggactac 420
agaaacgttc tgcaccgtga tgg 443
```

<210> 13  
 <211> 938  
 <212> DNA  
 <213> Arabidopsis thaliana  
  
 <220>  
 <221> unsure  
 <222> (1..938)  
 <223> unsure at all n locations  
  
 <400> 13

atgatactgc cagctannnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60  
 nnnnnnnnnn nnnnnnnnnn nnnccacgcg tccgaaaacg ttttatactg agtttctttc 120  
 accatccagc ttcatttggtg aaaaatcgtc aatccctctc aaactcttct caccactaat 180  
 ttcttctctt ggaacattct cttctctatt attttgattc ccttggcctc aacactgggtt 240  
 tctcaattgc atgatcttgg ctgctcttca gttactttga ttcactgaga aaaatggcga 300  
 ctggagtatt gccagctccg gtttctggga tcaagatacc ggattcgaaa gtcggggtttg 360  
 gtaaaagcat gaatcttggtg agaatttggtg atgtaggag tctaagatct gctaggagaa 420  
 gagtttcggt tatccggaat tcaaaccaag gctctgattt agctgagctt caacctgcat 480  
 ccgaaggaag ccctctctta gtgccaagac agaaatattg tgaatcattg cataagacgg 540  
 tgagaaggaa gactcgtact gttatggttg gaaatgtcgc ccttggaagc gaacatccga 600  
 taaggattca aacgatgact acttcggata caaaagatat tactggaact gttgatgagg 660  
 ttatgagaat agcggataaa ggagctgata ttgtaaggat aactgtccaa gggaagaaaag 720  
 aggcggatgc gtgctttgaa ataaaagata aactcgttca gcttaattac aatataccgc 780  
 tggttgacga tattcattgt gcccctactg tagccttacg agtcgctgaa tgctttgaca 840  
 agatccgtgt caaccagga aattttgcgg acaggcgggc ccagtttgag acgattgatt 900  
 atacagaaga tgaatatcag aaagaactcc agcatatc 938

<210> 14  
 <211> 432  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 14

agcataacaa ggctctgatt tagctgagct tcaacctgca tccgaaggaa gccctctctt 60  
 agtgccaaga cagaaatatt gtgaatcatt gcataagacg gtgagaagga agactcgtac 120

tgttatggtt ggaaatgtcg cccttggaag cgaacatccg ataaggattc aaacgatgac 180  
 tacttcggat acaaaagata ttactggaac tgttgatgag gttatgagaa tagcggataa 240  
 aggagctgat attgtaagga taactgttca agggaagaaa gaggcggatg cgtgctttga 300  
 aataaaagat aaactcgttc agcttaatta caatataaccg ctggttgcag atattcattt 360  
 tgcccctact gtagccttac gagtcgctga atgctttgac aagatccgtg tcaaccaag 420  
 aaattttgcg ga 432

<210> 15  
 <211> 528  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> unsure  
 <222> (1..528)  
 <223> unsure at all n locations  
 <400> 15

tgatacgcca gctctatacg actcactatt agggaagctg gtacgcctgc aggtacccgg 60  
 tccgggaatt ccnnggtcg acccacgcgt ccgaaagaac tccagcatat cgagcaggtc 120  
 ttcactcctt tggttgagaa atgcaaaaag tacgggagag caatgcgtat tgggacaaat 180  
 catggaagtc tttctgaccg tatcatgagc tattacgggg attctccccg aggaatggtt 240  
 gaatctgcgt ttgagtttgc aagaatatgt cggaaattag actatcacia ctttgttttc 300  
 tcaatgaaag cgagcaaccc agtgatcatg gtccaggcgt accgtttact tgtggctgag 360  
 atgtatgttc atggatggga ttatcctttg catttgggag ttactgaggc aggagaaggc 420  
 gaagatggac ggatgaaatc tgcgattgga attgggacgc ttcttcagga cgggctcggt 480  
 gacacaataa gagtttcact gacggagcca ccagaagagg agatagat 528

<210> 16  
 <211> 379  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <400> 16

gcgtattggg acaaatcatg gaagtctttc tgaccgtatc atgagctatt acggggattc 60  
 tccccgagga atggttgaat ctgcgtttga gtttgcaaga atatgtcgga aattagacta 120  
 tcacaacttt gttttctcaa tgaaagcgag caaccagtg atcatggtcc aggcgtaccg 180

tttactttgtg gctgagatgt atgttcatgg atgggattat cctttgcatt tgggagttac 240  
 tgaggcagga gaaggcgaag atggacggat gaaatctgcg attggaattg ggacgcttct 300  
 tcaggacggg ctcggtgaca caataagagt ttactgacg gagccaccag aagaggagat 360  
 agatccctgc aagcgattg 379

<210> 17  
 <211> 395  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 17

aaagaactcc agcatatcga gcaggtcttc actcctttgg ttgagaaatg caaaaagtac 60  
 gggagagcaa tgcgtattgg gacaaatcat ggaagtcttt ctgaccgtat catgagctat 120  
 tacggggatt ctccccgagg aatggttgaa tctgcgtttg agtttgcaag aatatgtcgg 180  
 aaattagact atcacaactt tgttttctca atgaaagcga gcaaccagat gatcatgggc 240  
 caggcgtacc gtttacttgt ggctgagatg tatgttcatg gatgggatta tcctttgcat 300  
 ttgggagtta ctgaggcagg agaaggcgaa gatggacgga tgaaatctgc gattggaatt 360  
 ggggacactt cttcaggacg ggctcggtga cacaa 395

<210> 18  
 <211> 395  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 18

aaagaactcc agcatatcga gcaggtcttc actcctttgg ttgagaaatg caaaaagtac 60  
 gggagagcaa tgcgtattgg gacaaatcat ggaagtcttt ctgaccgtat catgagctat 120  
 tacggggatt ctccccgagg aatggttgaa tctgcgtttg agtttgcaag aatatgtcgg 180  
 gaattagact atcacaactt tgttttctca atgaaagcga gcaaccagat gatcatgggc 240  
 caggcgtacc gtttacttgt ggctgagatg tatgttcatg gatgggatta tcctttgcat 300  
 ttgggagtta ctgatgcagg agaaggcgaa gatggacgga tgaaatctgc gattggaatt 360  
 gggacgcttc ttcaggacgg gctcggtgac acaat 395

<210> 19  
 <211> 412  
 <212> DNA  
 <213> Arabidopsis thaliana  
  
 <400> 19  
  
 atgctggagg ccttcttggt gatggactag gtgatggcgt aatgctcgaa gcacctgacc 60  
 aagattttga ttttcttagg aatacttcct tcaacttatt acaaggatgc agaatgcgta 120  
 acactaagac ggaatatgta tcgtgcccggt cttgtggaag aacgcttttc gacttgcaag 180  
 aaatcagcgc cgagatccga gaaaagactt cccatttacc tggcgtttcg atcgcaatca 240  
 tgggatgcat tgtgaatgga ccaggagaaa tggcagatgc tgatttcgga tatgtaggtg 300  
 gttctcccgg aaaaatcgac ctttatgtcg gaaagacggt ggtgaagcgt gggatagcta 360  
 tgacggaggc aacagatgct ctgatcggtc tgatcaaaga acatggtcgt tg 412

<210> 20  
 <211> 1172  
 <212> DNA  
 <213> Arabidopsis thaliana  
  
 <220>  
 <221> unsure  
 <222> (1..1172)  
 <223> unsure at all n locations  
  
 <400> 20  
  
 gggatatgcca ttcaaggatc tggcaactgt tgattcaatc ttattaaaga gagctaccgc 60  
 ctgtagatga tcaagtggct cgtttggtc taaaacgggt gattgatgtc agtatgggag 120  
 ttatagcacc tttatcagag caactaaca agccattgcc caatgccatg gttcttggtca 180  
 acctcaagga actatctggt ggcgcttaca agcttctccc tgaaggatca cgcttggttg 240  
 tctctctacg aggcgatgag ccttacgagg agcttgaaat actcaacaac attgatgcta 300  
 cgatgattct ccatgatgta cctttcactg aagacaaagt tagcagagta catgcagctc 360  
 ggaggctatt cgagttctta tccgagaatt cagttaactt tcctgttatt catcacataa 420  
 acttcccaac cggaatccac agagacgaat tggatgattca tgcagggaca tatgctggag 480  
 gccttcttgt ggatggacta cgtgatggcg taatgctcga agcacctgac caagattttg 540  
 attttcttag gaatacttcc ttcaacttat tacaaggatg cagaatgcgt aacactaaga 600  
 cggaatatgt atcgtgcccg tcttgtggaa gaacgctttt cgacttgcaa gaaatcagcg 660  
 ccgagatccg agaaaagact tcccatctac ctggcgtttc gatcgcaatc atgggatgca 720

ttgtgaatgg accaggagaa atggcagatg ctgatttcgg atatgtaggt gggtctcccg 780  
 gaaaaatcga cctttatgtc ggaaagacgg tggatgaagcg tgggatagct atgacggagg 840  
 caacagatgc tctgatcggg ctgatcaaag aacatgggctg ttgggtcgac ccgcccgtgg 900  
 ccgatgagta gatttcaaaa cggagaaaaga tgggtgggcc attctttgaa aactgtgaga 960  
 ggagatatat atatttgtgt gtgtatatca tctgtttgtt gtgtattgca tcattcattt 1020  
 tggacaaatg tccaaattct cttaagttga taaaagttct taggccaaat taaatttaat 1080  
 ataaaaaaaa aaaaaaaaaa gcnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140  
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nn 1172

<210> 21  
 <211> 584  
 <212> DNA  
 <213> Zea mays

<400> 21  
 caggtaatt aattcctgta cgccgtcggg ttccgggtact cgtttaattt cttcccgacc 60  
 acggttgatg gcaatgtaac cggcttggtt accacatag ccatagtcgg catcggccat 120  
 ttccccgggg ccattgacaa tacagcccat gacggcgatg tctaaaccg ttagatgttt 180  
 agtggcttct cggacttcat gtaacacgtc ttccaagttg aacaacgtgc ggccacagga 240  
 aggacaggcc acatattcca ccatggtttt ccgcaaacc agcgcttgga gaatgctgta 300  
 gcaaacggga atttcttttt cgggggcttc ggtgagggat acccggatag tatcgccaat 360  
 gccatcagct aaaagggtgg caatgccagc ggtggattta atgcggccat attccccatc 420  
 cccggcttcg gtaacccta gatggagggg ataatccatg cccaactcgt tcatacgttt 480  
 caccatgagg cgataggcgg ccaacattac cggtagccgg gacgctttca tggaaacgac 540  
 taggttgagg aaatctaaag actcacaat tttgatgaat tcca 584

<210> 22  
 <211> 670  
 <212> DNA  
 <213> Zea mays

<400> 22  
 caggtcgact ctagaggatc ggcgttaacc atggttctct ctccgaaaga atgcttttac 60  
 ctacttttta cccccgaggg catggtgcaa tcggccctgg aattcatcaa aatttgtgag 120



tccttagatt tccgcaacct agtcgtttcc atgaaagcgt cccgggtacc ggtaatgttg 180  
gccgcctatc gcctcatggt gaaacgtatg gacgagttgg gcatggatta tcccctccat 240  
ctaggggtta ccgaagccgg ggatggggaa tatggccgca ttaaateccac cgctggcatt 300  
gccacccttt tagctgatgg cattggcgat actatccggg tatccctcac cgaagcccc 360  
gaaaaagaaa ttcccgtttg ctacagcatt ctccaggcgc tgggtttgcg gaaaaccatg 420  
gtggaatatg tggcctgtcc ttctgtggc cgcacgttgt tcaacttgga agacgtgtta 480  
catgaagtcc gagatgccac taaacatcta acgggttttag actttcgccg tcatgggctg 540  
tattgtcaat ggccccgggg caatggccga tgcgactat ggctatgtgg gtaaacaagc 600  
cggttacatt gccatcaacc gtggtcggga agaaattaaa cgagtaccg aaaccgacgg 660  
cgtacaggaa 670

<210> 23  
<211> 596  
<212> DNA  
<213> Zea mays  
<220>  
<221> unsure  
<222> (1..596)  
<223> unsure at all n locations  
<400> 23

caggctgact ctagaggatc ggcgttaacc atggttctct ctccgaaaga atgcttttac 60  
ctacttttta ccccgaggg catggtgcaa tcggccctgg aattcatcaa aatttgtgag 120  
tccttagatt tccgcaacct agtcgtttcc atgaaagcgt cccgggtacc ggtaatgttg 180  
gccgcctatc gcctcatggt gaaacgtatg gacgagttgg gcatggatta tcccctccat 240  
ctaggggtta ccgaagccgg ggatggggaa tatggccgca ttaaateccac cgctggcatt 300  
gccacccttt tagctgatgg cattggcgat actatccggg tatccctcac cgaagcccc 360  
gaaaaagaaa ttcccgtttg ctacagcatt ctccaggcgc tgggtttgcg gaaaaccatg 420  
gtggaatatg tggcctgtcc ttctgtggc cgcacgttgt tcaacttgga agacgtgtta 480  
catgaagtcc gagatgccac taaacatcta acgtgttttag actttcgncg tcatgtgctg 540  
tattgtcaat ggccccgggtg caatggccga tgcgactat ggctatgtgg gtaaac 596

<210> 24  
 <211> 403  
 <212> DNA  
 <213> Zea mays  
  
 <400> 24  
  
 cagacaagga ggaggaaaac tcgaactgtg atggtgggga atgtgccact tgggagtgat 60  
 caccataa ggattcaaac catgacgact tcagatacca aggatgttgc gaaaacagta 120  
 gaggaggtga tgaggatagc agataaagga gctgatcttg ttagaataac agtccagggt 180  
 aggaaggaag ctgatgcctg ctttgagatc aagaacactc tggttcagaa gaattacaac 240  
 attccactag tggccgatat tcattttgct cctacggtag ctctaaagggt ggcagaatgt 300  
 tttgacaaaa ttcgtgtgaa cccaggaaat tttgctgac gtcgtgctca atttgaaaag 360  
 ctggaatata ctgacgacga ctacaaaaa gagctagagc ata 403

25  
 293  
 DNA  
 Zea mays  
 25  
 60  
 120  
 180  
 240  
 293

<210> 25  
 <211> 293  
 <212> DNA  
 <213> Zea mays  
  
 <400> 25  
  
 cagacaaggc ggaggaaaac tcgaactgtg atggtgggga atgtgccact tggcagtgat 60  
 caccataa ggattcaaac catgacgact tcagatacca aggatgttgc gaaaacagta 120  
 gaggaggtga tgaggatagc agataaagga gctgatcttg ttagaataac agtccagggt 180  
 aggaaggaag ctgatgcctg ctttgagatc aagaacactc tggttcagaa gaattacaac 240  
 attccactag tggccgatat tcattttgct cctacggtag ctctaagggt ggc 293

<210> 26  
 <211> 456  
 <212> DNA  
 <213> Zea mays  
  
 <400> 26  
  
 cagacaaggc ggaggaaaac tcgaactgtg atggtgggga atgtgccact tggcagtgat 60  
 caccataa ggattcaaac catgacgact tcagatacca aggatgttgc gaaaacagta 120  
 gaggaggtga tgaggattgc agataaagga gctgatcttg ttagaataac agtccagggt 180  
 aggaaggaag ctgatgcctg ctttgagatc aagaacaact ctggttcaga agaattacaa 240  
 ccttccacta gtggacctga tattcatttt gtccttcag tagctttaa ggtggcagaa 300

tgtttggaca aattaattga aacacacaat ttcttgttga tagtgtacct taattagaaa 360  
 agctggaatt taccggctac gacttccata aagcgcttgg gcttgtttaa caattgggtt 420  
 ttaccttaat cgaatatttc acagaaattt gaattt 456

<210> 27  
 <211> 619  
 <212> DNA  
 <213> Zea mays

<400> 27

caccgaaggt ttctaattta tttctcagat ctcaataaat gtacaaaatg tgtaggggatg 60  
 atgtacattg tatgctcagt tcctgcattg cgtgtttcgc ttacagaat atataaacta 120  
 cagacttggc tacagcctac agccctactc ctcggcagga ggatccaccc atcggccatg 180  
 gtccttgatc agctggatca aggcgtcagt tgcaccttcc atggcgatgg cgcgctgcac 240  
 aacggtcttg ccaacataaa ggtcgatctt tccgggagcg cctccaacgt atccgaaatc 300  
 ggcattcagc atctctcctg gtccattgac aatacaaccc atgatatgca tcgaaacacc 360  
 tggcagatga gaggtctttt ctctaatttc agcgtgatt tcctgaaggt caaagagtgt 420  
 tcggccgcag gaaggacaag acacatatc agtttttgtg ttgcgcattc tgcaaccttg 480  
 gagcaagttg aaagatgtgt ccctcaggaa ctcaaattcc tggtcagcag cttcaaggaa 540  
 tacaccatca ccaagaccat cgactaagag agcaccaacg ttggccccag caccaatgac 600  
 aagaccatct ctgtcaatg 619

<210> 28  
 <211> 422  
 <212> DNA  
 <213> Zea mays

<400> 28

tcgcttgac ttgggtgtta cagaagctgg agagggtgaa gatggaagga tgaaatctgc 60  
 tattggcatt gggacactgc taatggatgg tttgggtgat acaatccgtg tctccctcac 120  
 agaaccacca gaagaagaga ttgatccttg ccaaagggtg gcaaattctg ggacgcaggc 180  
 cgcaaacctt caaattgggg tggccccatt tgaagaaaag cacaggcgct attttgattt 240  
 ccagcgtagg agtgggtcaat tgcccttgca gaaggaggga ggcgatagtt gactacagaa 300  
 atgtcctgca tcgtgatggg atctgactga tggcagtttc cctggatcag ttgaaggctc 360

ctgatctcct ttataggtat attgcagcaa agcttgcgga tggcatgcct ttcaaggatc 420

tg 422

<210> 29  
<211> 430  
<212> DNA  
<213> Zea mays

<400> 29

tcgcttgacac ttgggtgtta cagaagctgg agagggtgaa gatggaagga tgaaatctgc 60

tattggcatt gggacactgc taatggatgg tttgggtgat acaatccgtg tctccctcac 120

agaaccacca gaagaagaga ttgataccttg ccaaagggtg gcaaatacttg ggacgcaggc 180

tgcaaacctt caaattgggg tggccccatt tgaagaaaag cacaggcggt attttgattt 240

ccagcgtagg agtgggtcaat tgcctttgca gaaggagggt gaggaagttg actacagaaa 300

tgtcctgcat cgtgatggta tctgtactga tggcagtttc cctggatcag ttgaaggctc 360

ctgatctcct ttataggtct cttgcagcaa agcttgcggt tggcatgcct ttcaaggatc 420

tggctactgt 430

<210> 30  
<211> 528  
<212> DNA  
<213> Zea mays

<400> 30

gacaggcagg gtgcatgctg ctaggaggtt atttgagtac ttacaggcca atggcttgaa 60

cttccctgta attcatcaca taaatttccc tgaaaccatt gacagagatg gtcttgtcat 120

tggggctggg gccaacgttg gtgctctctt agtcgatggg cttggtgatg gtgtattcct 180

tgaggcggct gaccaggaat ttgagttcct gagggacaca tctttcaact tgctccaagg 240

ttgcaggatg cgcaacacaa aaactgaata tgtgtcttgt ccttcctgcg gccgaacact 300

ctttgacctt caggaaatca gcgctgagat tagcgaaaag acctctcatc tgccacgtgt 360

ttcgatcgct atcatgggtt gtattgtcaa tggaccagga gcgctggctg atgccgattt 420

cggatacgtt ggcggcgctc ccggaaagat cgacctttat attggcacga ccgttatgca 480

gcgcgccatc gccatggacg gtgcaactga cgccttgatc cagctgat 528

<210> 31  
 <211> 303  
 <212> DNA  
 <213> Zea mays

<400> 31

```

ggggccaacg ttggtgctct cttagtcgat ggtcttggtg atggtgtatt ccttgaggcg 60
gctgaccagg aatttgagtt cctgagggac acatctttca acttgctcca aggttgcagg 120
atgcgcaaca caaaaactga atatgtgtct tgtccttcct gcggccgaac actctttgac 180
cttcaggaaa tcagcgctga gattagagaa aagacctctc atctgccacg tgtttcgatc 240
gctatcatgg gttgtattgt caatggacca ggagagatgg ctgatgccga tttcggatac 300
gtt 303
  
```

<210> 32  
 <211> 613  
 <212> DNA  
 <213> Zea mays

<220>  
 <221> unsure  
 <222> (1..613)  
 <223> unsure at all n locations

<400> 32

```

cgagatggcg ttccatgcn ggcccttcct cctcttcctc ttctttctgcc cccccgctgg 60
cttgaaaaag ggagagaaac tcgcgcactc ggttatcgaa gggaggagcg cgggcgaggg 120
tgaggtttcg cccacacgga gctgcgaggt gtttgtagga tctcctaggt gagcccctgc 180
tgcttgagga cagccatggc caccggcgtg gctccagctc ctctcccaca tgtcagagtg 240
cgtcatgggg gcgtcggggt caccaggagc gtcgattttg cgaagggtctt gtctgctccc 300
ggtgccggca cgatgagagc aagctcctct agaggcaggg cgctcgtggc gaagagctct 360
agtactggct cggagaccat ggagctcgag ccatcttcag aaggaagccc acttttagta 420
cccaggcaga agtactgtga atcaacacac cagacaagga ggaggaaaac tcgaactgtg 480
atggtgggga atgtgccact tggcagtgat catcccataa ggattcaaac catgacgact 540
tcagatacca aggatgttgc aaaaacagta gaggaggtga tgaggatagc agataaagga 600
gctgatcttg tta 613
  
```

<210> 33  
 <211> 464  
 <212> DNA  
 <213> Glycine max

<400> 33

agagcatgaa atcttctgcg aggaaaaggg tgtcaattat cacgaactca aatcctggcc 60  
 aagatattgc tgaacttcaa cctgcatccc caggaagccc tcttttggtt cctaggcaaa 120  
 agtattgtga atcattgcac aaacccatca ggagaaaaac aagcacagta atggttggtgta 180  
 acgtggctat tggtagcgag catcctataa gaattcagac catgactaca actgacacta 240  
 aggatgttgc tgggacagtt gaacaggtga tgagaatagc agataaagga gctgatattg 300  
 tacggataac agttcaaggg aagaaagaag ctgatgcttg ttttgagatt aaaaacaccc 360  
 ttgtgcagaa aaattacaac ataccctggg tggctgatat tcattttgct ccctctgttg 420  
 ctttgcgggt agctgaatgc tttgataaga ttcgtgtaaa ccct 464

<210> 34  
 <211> 705  
 <212> DNA  
 <213> Glycine max

<400> 34

gtagctgaat gctttgataa gattcgtgta aaccctggaa attttgctga tagacgggct 60  
 caatttgaac cattagagta cacagaagaa gactatcaga aagaacttga gcatattgaa 120  
 aaggttttca caccattggt tgagaaatgt aagaaatatg ggagagcaat gcgcattggg 180  
 acaaaccatg gaagtctttc tgatcgtata atgagctact atggagactc gcctagggga 240  
 atggtagaat ctgcttttga atttgcaagg atatgccgaa agttagacta tcacaatttt 300  
 gttttttcta tgaaagcaag caaccagtt atcatggttc aggcataaccg cttacttgtg 360  
 gctgaaatgt atgtccaagg ctgggattat ccattacact tgggtgttac tgaagctgga 420  
 gaagggtgagg atgggaggat gaagtctgca ataggcattg gaactcttct tcaggatgga 480  
 ttgggtgata caattagggt ttctctcaca gaaccaccag aggaggagat agacccttgc 540  
 agaaggttgg caaatcttgg aatgatagct tctgaactcc agaagggggg ggaacctttt 600  
 gaagaaaagc acagacatta ttttcgactt tcagcgccga tctggtcaat tgccagtgca 660  
 aaaagagggt gaggaggtgg attacagagg tgtactgcac cgtga 705

<210> 35  
 <211> 564  
 <212> DNA  
 <213> Glycine max

<220>  
 <221> unsure  
 <222> (1..564)  
 <223> unsure at all n locations

<400> 35

```
aagcncggaa ttcggctcga gaggaactca aatcctggcc aagatattgc tgaacttcaa 60
cctgtatccc caggaagccc tcttttggtt cctaggcaaa agtattgtga atgattacac 120
aaaactgtca ggagaaaaac aaacacagtg atggttggtg acgtggctat tggtagcgag 180
catcctataa gaattcagac catgactacg actgacacta aggatgttgc tgggacagtt 240
gaacagggtg tgagaatagc agataaagga gctgatattg tacggataac agttcaaggg 300
aagaaagaag ctgatgcttg ttttgagatt aaaaacaccc ttgttcagaa aaattacaac 360
atactcgtgg tggctgatat tcattttgct ccctctggtg ctttgcggtt agctgaatgc 420
tttgataaga ttcgtgtaaa ccctggaaat tttgctgata gacgggctca atttgaaaca 480
ttagagtaca cagatgatga ctatcagaaa gaacttgagc atattgaaaa ggttttcaca 540
ccattggttg agaaatgtaa gaaa 564
```

<210> 36  
 <211> 511  
 <212> DNA  
 <213> Glycine max  
 <400> 36

```
aaaccatgga agtctttctg atcgtataat gagctactat ggagactcgc ctagggggaat 60
ggtagaatct gcttttgaat ttgcaaggat atgccgaaag ttagactatc acaattttgt 120
tttttctatg aaagcaagca acccagttat catggttcag gcataccgct tacttgtggc 180
tgaaatgtat gtccaaggct gggattatcc attacacttg ggtgttactg aagctggaga 240
aggtgaggat gggaggatga agtctgcaat aggcattgga actcttcttc aggatggatt 300
gggtgataca attaggggtt ctctcacaga accaccagag gaggagatag acccttgcag 360
aaggttggca aatcttggaa tgatagcttc tgaactccag aaggggggtg aaccttttga 420
agaaaagcac agacattatt ttgactttca gcgccgatct ggtcaattgc cagtgcataa 480
agagggtgag gaggtggatt acagaggtgt a 511
```

<210> 37  
 <211> 498  
 <212> DNA  
 <213> Glycine max

<220>  
 <221> unsure  
 <222> (1..498)  
 <223> unsure at all n locations

<400> 37

cggaggtggc gtgaatgctt tgataagatt cgtgtaaacc ctggaaattt tgctgataga 60  
 cgggctcaat ttgaaacatg agagtggaca naataagact atgagaaaga acttgagcat 120  
 attgaaaagg ttttcacacc attggttgag aaatgtaaga aatatgggag agcaatgcgc 180  
 attgggacaa accatggaag tctttctgat cgtataatga gctactatgg agactcgcct 240  
 aggggaatgg tagaatctgc ttttgaattt gcaaggatat gccgaaagtt agactatcac 300  
 aattttgttt tttctatgaa agcaagcaac ccagttatca tggttcaggc ataccgctta 360  
 cttgtggctg aaatgtatgt ccaaggctgg gattatccat tacacttggg tgttactgaa 420  
 gctggagaag gtgaggatgg gaggatgaag tctgcaatag gcattggaac tcttcttcag 480  
 gatggattgg gtgataca 498

<210> 38  
 <211> 440  
 <212> DNA  
 <213> Glycine max  
 <400> 38

gtagctgaat gctttgataa gattcgtgta aaccctggaa attttgttga tagacgggct 60  
 caatttgaaa cattagagta cacagaagaa gactatcata aagaacttga gcatattgaa 120  
 aaggttttca caccattggg tgagaaatgt aagaaatatg ggagagcaat gcgcattggg 180  
 acaaaccatg gaagtctttc tgatcgtata atgagctact atggagactc gcctagggga 240  
 atggtagaat ctgcttttga atttgcaagg atatgccgaa agttagacta tcacaatttt 300  
 gttttttcta tgaaagcaag caaccagtt atcatggttc aggcataccg cttacttgtg 360  
 gctgaaatgt atgttcaagg ctgggattat ccattacact tgggtgttac tgaagctgga 420  
 aaaagtgagg atgggaggat 440



<210> 39  
 <211> 353  
 <212> DNA  
 <213> Glycine max

<400> 39

aattcggctc gagaggaact caaatcctgg ccaagatatt gctgaacttc aacctgcac 60  
 cccaggaagc cctcttttgg ttcctaggca aaagtattgt gaatcattac acaaaactgt 120  
 caggagaaaa acaaacacag tgatgggttg taacgtggct attggtagcg agcatcctat 180  
 aagaattcag accatgacta cgactgacac taaggatggt gctgggacag ttgaacaggt 240  
 gatgagaata gcagataaag gagctgatat tgtacggata acagttcaag ggaagaaaga 300  
 agctgatgct tgttttgaga ttaaaaacac ccttgttcaa aaaaattaca aca 353

<210> 40  
 <211> 577  
 <212> DNA  
 <213> Glycine max

<400> 40

gatgtttttg tcgtgtattc tttcctatt gcattcagct cactgatttc aattacaaag 60  
 tcaattttgt aaatcagagg cagagagagt tgtaaagagc ctctgaattt tgatcacacc 120  
 acacccttct tctcatctcc accagaaatg gctaccggag ctgctgtgcc aactacgttt 180  
 tctaccctca agacatggga ttccagtttg gggtttgcaa aaaacataga ttttgtgaga 240  
 gtttccgata tgaagagcat gaaatcttct gcgaggaaaa ggggtgtcaat ttcaggaac 300  
 tcaaatcctg gccagatat tgctgaactt caacctgcat cccaggaag ccctcttttg 360  
 gttcctaggc aaaagtattg tgaatcattg cacaacccta tcaggagaaa aacaagcaca 420  
 gtaatggttg gtaacgtggc tattggtagc gagcatccta taagaattca gaccatgact 480  
 acaactgaca ctaaggatgt tgctgggaca gttgaaccgg tgatgagaat agcagataaa 540  
 ggagctgata ttgtacggat aacagttcaa gggaaga 577

<210> 41  
 <211> 551  
 <212> DNA  
 <213> Glycine max

<400> 41

tggtgctggg tctgatgctg gagcccttct ggtggatggg cttggagatg gacttctttt 60

ggaagcgcca gacaaggatt ttgaatttat tagaaacact tctttcaatt tgttgcaagg 120  
 ctgcagaatg agaaatacaa agacagagta tgtctcatgt ccatcctgtg gcagaacatt 180  
 gtttgatctt caagaagtaa gtgcacaaat tcgggagaag acatcacacc tccccggtgt 240  
 ttcgattgca atcatgggat gcattgtaaa tggaccaggg gagatggctg atgcagactt 300  
 tgggtatgtg ggaggcactc ccgggaagat tgacctctat gttgggaaga ctgtggtgaa 360  
 gcgtggaatt gcaatggagc atgcaaccaa tgccttgatc gatctaataa aagaacatgg 420  
 acgatgggtg gaccctcctg ccgaggagta aaagcaagag cttaatTTTg agattggcat 480  
 tcaaggccat agtaagatga gcattgtcat atccaattat tggacacatg taatataagc 540  
 atacactcaa t 551

<210> 42  
 <211> 869  
 <212> DNA  
 <213> Glycine max  
 <400> 42

gaagcatagt agcatcaatg ccttccttat acagaagact aaaattagca gagtgcattgc 60  
 ggccaggcgg ttatttgagt acctatccga caattctcta aacttccttg ttattcacca 120  
 tattcagttc ccaaattggga ttcacagaga tgacttggtg attggtgctg gttctgatgc 180  
 tggagccctt ctggtggatg ggcttggaga tggacttctt ttggaagcgc cagacaagga 240  
 ttttgaattt attagaaaca cttctttcaa tttgttgcaa ggctgcagaa tgagaaatac 300  
 aaagacagag tatgtctcat gtccatcctg tggcagaaca ttgtttgatc ttcaagaagt 360  
 aagtgcacaa attcgggaga agacatcaca cctccctggg gtttcgattg caatcatggg 420  
 atgcattgta aatggaccag gggagatggc tgatgcagac tttgggtatg tgggaggcac 480  
 tccccggaag attgacctct atgttgggaa gactgtggtg aagcgtggaa ttgcaatgga 540  
 gcatgcaacc aatgccttga tcgatctaataaaaagaacat ggacgatggg tggaccctcc 600  
 tgccgaggag taaaagcaag agcttaattt tgagattggc attcaaggcc atagtaagat 660  
 gagcattgtc atatccaatt attgtacaca tgtaataataa gataacactc aatgcttaag 720  
 tttgagccta gttttaagtt ccttttgaga aagatcccaa ttaaagcttg ttgtgaggaa 780  
 atcgacagct agaacatgta tacagataac agtgtattgc tttgccccat cagccatcaa 840  
 taataatgag aatctcttag aatagtgcc 869

<210> 43  
 <211> 291  
 <212> DNA  
 <213> Glycine max

<220>  
 <221> unsure  
 <222> (1..291)  
 <223> unsure at all n locations

<400> 43

gangnactca aatcctgggc caagatattg ctgaacttca nccctgcac cccaggngc 60  
 cctcttttgg ttcctaggca aaagtattgt gaatcattnc cacaaaactg nccagganaa 120  
 aaacaaacac agtgatgggt ggtaacgtgg ctattggtag cgagcatcct ataagaattc 180  
 agaccatgac tacgacngac actaaggatg ttgctgggac agtngaacng gtgatgagaa 240  
 tagcagataa aggagctgat attgtacgga taacagttca agggaagaaa g 291

<210> 44  
 <211> 388  
 <212> DNA  
 <213> Glycine max

<400> 44

cccggtatat gggtcaggca taccgtttac ttgtggctga aatgtatgtc caaggctggg 60  
 attatccatt acacttgggt gttactgaag ctggagaagg tgaggatggg aggatgaagt 120  
 ctgcaattgg cattggaact cttcttcagg atggattggg tgatacaatt agggtttctc 180  
 tcacagaacc accagaagag gagatagatc cttgcagaag gttggcaaat cttggaatga 240  
 gagcttctga actccagaag ggggtggaac cttttgaaga aaagcacaga cattattttg 300  
 acttccagcg ccgatctggt caattgccag tgcaaaaaga gggtgaggag gtggattaca 360  
 gaggtgcact gcaccgtgac gggttctgt 388

<210> 45  
 <211> 211  
 <212> DNA  
 <213> Glycine max

<400> 45

cccggttata atggcgagg cataccgctt acttgtggct gaaatgtatg tccaaggctg 60  
 ggattatcca ttacacttgg gtgttactga agctggagga ggtgaggatg acaggatgaa 120

gtctgcaatt ggcattggaa ctcttcttca ggatggattg ggtgatacaa ttaggggtgtc 180  
tcgcacagaa ccaccagaag aggagataga t 211

<210> 46  
<211> 276  
<212> DNA  
<213> Glycine max  
<400> 46

tgggcttgga gatggactac ttttggaagc cccggacaag gattttgaat ttattagaaa 60  
cacttctttc aatttggtgc aaggctgcag aatgagaaat acaaagacag agtatgtctc 120  
atgtccatcc tgtggcagaa cattgtttga tcttcaagaa gtaagtgcac aaattcggga 180  
gaagacatca cacctccctg gtgtttcgat tgcaatcatg ggatgcattg taaatggacc 240  
aggggagatg gctgatgcag actttgggta tgtggg 276

CCAGGAGATG GCTGATGCAG ACTTTGGGTA TGTGGG

<210> 47  
<211> 399  
<212> DNA  
<213> Brassica napus  
<400> 47

cccacgcgtc cgcagggatt cacagggacg agttggtgat ccacgcaggg acatacgtctg 60  
gggcacttct agtggatgga cttggagatg gtgtaatgct agaagcacct gatcaagact 120  
tcgagtttct taggaacact tctttcaact tgttacaagg ctgcaggatg cgtaacacca 180  
agacggaata cgtatcgtgc ccgtcttggtg gaagaactct gttcgacttg caagaaatca 240  
gcgctgagat cagagaaaag acttcgcatt tgccctggcgt ttcgattgca ataatggggtt 300  
gcattgtgaa tggacctggc gaaatggctg atgctgattt cggttatgta ggcggttctc 360  
ccgggaaaat cgacctttac gttggaaaga cgggtggtca 399

<210> 48  
<211> 740  
<212> PRT  
<213> Arabidopsis thaliana  
<400> 48

Met Ala Thr Gly Val Leu Pro Ala Pro Val Ser Gly Ile Lys Ile Pro  
1 5 10 15

Asp Ser Lys Val Gly Phe Gly Lys Ser Met Asn Leu Val Arg Ile Cys  
20 25 30

Asp	Val	Arg	Ser	Leu	Arg	Ser	Ala	Arg	Arg	Arg	Val	Ser	Val	Ile	Arg	
	35						40					45				
Asn	Ser	Asn	Gln	Gly	Ser	Asp	Leu	Ala	Glu	Leu	Gln	Pro	Ala	Ser	Glu	
	50					55					60					
Gly	Ser	Pro	Leu	Leu	Val	Pro	Arg	Gln	Lys	Tyr	Cys	Glu	Ser	Leu	His	
	65				70					75					80	
Lys	Thr	Val	Arg	Arg	Lys	Thr	Arg	Thr	Val	Met	Val	Gly	Asn	Val	Ala	
				85					90					95		
Leu	Gly	Ser	Glu	His	Pro	Ile	Arg	Ile	Gln	Thr	Met	Thr	Thr	Ser	Asp	
			100					105						110		
Thr	Lys	Asp	Ile	Thr	Gly	Thr	Val	Asp	Glu	Val	Met	Arg	Ile	Ala	Asp	
		115					120					125				
Lys	Gly	Ala	Asp	Ile	Val	Arg	Ile	Thr	Val	Gln	Gly	Lys	Lys	Glu	Ala	
	130					135					140					
Asp	Ala	Cys	Phe	Glu	Ile	Lys	Asp	Lys	Leu	Val	Gln	Leu	Asn	Tyr	Asn	
	145				150					155					160	
Ile	Pro	Leu	Val	Ala	Asp	Ile	His	Phe	Ala	Pro	Thr	Val	Ala	Leu	Arg	
				165					170					175		
Val	Ala	Glu	Cys	Phe	Asp	Lys	Ile	Arg	Val	Asn	Pro	Gly	Asn	Phe	Ala	
			180					185					190			
Asp	Arg	Arg	Ala	Gln	Phe	Glu	Thr	Ile	Asp	Tyr	Thr	Glu	Asp	Glu	Tyr	
		195					200					205				
Gln	Lys	Glu	Leu	Gln	His	Ile	Glu	Gln	Val	Phe	Thr	Pro	Leu	Val	Glu	
	210					215					220					
Lys	Cys	Lys	Lys	Tyr	Gly	Arg	Ala	Met	Arg	Ile	Gly	Thr	Asn	His	Gly	
	225				230					235					240	
Ser	Leu	Ser	Asp	Arg	Ile	Met	Ser	Tyr	Tyr	Gly	Asp	Ser	Pro	Arg	Gly	
				245					250					255		
Met	Val	Glu	Ser	Ala	Phe	Glu	Phe	Ala	Arg	Ile	Cys	Arg	Lys	Leu	Asp	
			260					265					270			
Tyr	His	Asn	Phe	Val	Phe	Ser	Met	Lys	Ala	Ser	Asn	Pro	Val	Ile	Met	
		275					280					285				
Val	Gln	Ala	Tyr	Arg	Leu	Leu	Val	Ala	Glu	Met	Tyr	Val	His	Gly	Trp	
	290					295					300					
Asp	Tyr	Pro	Leu	His	Leu	Gly	Val	Thr	Glu	Ala	Gly	Glu	Gly	Glu	Asp	
	305				310					315					320	
Gly	Arg	Met	Lys	Ser	Ala	Ile	Gly	Ile	Gly	Thr	Leu	Leu	Gln	Asp	Gly	
				325					330					335		

Leu Gly Asp Thr Ile Arg Val Ser Leu Thr Glu Pro Pro Glu Glu Glu  
340 345 350

Ile Asp Pro Cys Arg Arg Leu Ala Asn Leu Gly Thr Lys Ala Ala Lys  
355 360 365

Leu Gln Gln Gly Ala Pro Phe Glu Glu Lys His Arg His Tyr Phe Asp  
370 375 380

Phe Gln Arg Arg Thr Gly Asp Leu Pro Val Gln Lys Glu Gly Glu Glu  
385 390 395 400

Val Asp Tyr Arg Asn Val Leu His Arg Asp Gly Ser Val Leu Met Ser  
405 410 415

Ile Ser Leu Asp Gln Leu Lys Ala Pro Glu Leu Leu Tyr Arg Ser Leu  
420 425 430

Ala Thr Lys Leu Val Val Gly Met Pro Phe Lys Asp Leu Ala Thr Val  
435 440 445

Asp Ser Ile Leu Leu Arg Glu Leu Pro Pro Val Asp Asp Gln Val Ala  
450 455 460

Arg Leu Ala Leu Lys Arg Leu Ile Asp Val Ser Met Gly Val Ile Ala  
465 470 475 480

Pro Leu Ser Glu Gln Leu Thr Lys Pro Leu Pro Asn Ala Met Val Leu  
485 490 495

Val Asn Leu Lys Glu Leu Ser Gly Gly Ala Tyr Lys Leu Leu Pro Glu  
500 505 510

Gly Thr Arg Leu Val Val Ser Leu Arg Gly Asp Glu Pro Tyr Glu Glu  
515 520 525

Leu Glu Ile Leu Lys Asn Ile Asp Ala Thr Met Ile Leu His Asp Val  
530 535 540

Pro Phe Thr Glu Asp Lys Val Ser Arg Val His Ala Ala Arg Arg Leu  
545 550 555 560

Phe Glu Phe Leu Ser Glu Asn Ser Val Asn Phe Pro Val Ile His His  
565 570 575

Ile Asn Phe Pro Thr Gly Ile His Arg Asp Glu Leu Val Ile His Ala  
580 585 590

Gly Thr Tyr Ala Gly Gly Leu Leu Val Asp Gly Leu Gly Asp Gly Val  
595 600 605

Met Leu Glu Ala Pro Asp Gln Asp Phe Asp Phe Leu Arg Asn Thr Ser  
610 615 620

Phe Asn Leu Leu Gln Gly Cys Arg Met Arg Asn Thr Lys Thr Glu Tyr  
625 630 635 640

Val Ser Cys Pro Ser Cys Gly Arg Thr Leu Phe Asp Leu Gln Glu Ile  
                     645                    650                    655  
 Ser Ala Glu Ile Arg Glu Lys Thr Ser His Leu Pro Gly Val Ser Ile  
                     660                    665                    670  
 Ala Ile Met Gly Cys Ile Val Asn Gly Pro Gly Glu Met Ala Asp Ala  
                     675                    680                    685  
 Asp Phe Gly Tyr Val Gly Gly Ser Pro Gly Lys Ile Asp Leu Tyr Val  
                     690                    695                    700  
 Gly Lys Thr Val Val Lys Arg Gly Ile Ala Met Thr Glu Ala Thr Asp  
                     705                    710                    715                    720  
 Ala Leu Ile Gly Leu Ile Lys Glu His Gly Arg Trp Val Asp Pro Pro  
                     725                    730                    735  
 Val Ala Asp Glu  
                     740

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

<210> 49  
 <211> 603  
 <212> PRT  
 <213> Oryza sativa  
 <400> 49  
 Met Val Gly Asn Val Pro Leu Gly Ser Asp His Pro Ile Arg Ile Gln  
 1                    5                    10                    15  
 Thr Met Thr Thr Ser Asp Thr Lys Asp Val Ala Lys Thr Val Glu Glu  
                     20                    25                    30  
 Val Met Arg Ile Ala Asp Lys Gly Ala Asp Phe Val Arg Ile Thr Val  
                     35                    40                    45  
 Gln Gly Arg Lys Glu Ala Asp Ala Cys Phe Glu Ile Lys Asn Thr Leu  
                     50                    55                    60  
 Val Gln Lys Asn Tyr Asn Ile Pro Leu Val Ala Asp Ile His Phe Ala  
 65                    70                    75                    80  
 Pro Thr Val Ala Leu Arg Val Ala Glu Cys Phe Asp Lys Ile Arg Val  
                     85                    90                    95  
 Asn Pro Gly Asn Phe Ala Asp Arg Arg Ala Gln Phe Glu Gln Leu Glu  
                     100                    105                    110  
 Tyr Thr Glu Asp Asp Tyr Gln Lys Glu Leu Glu His Ile Glu Lys Val  
                     115                    120                    125  
 Pro Asn Ile Ser Leu Phe Ser Val Asn Leu Val Phe Ser Pro Leu Val  
                     130                    135                    140

Glu Lys Cys Lys Gln Tyr Gly Arg Ala Met Arg Ile Gly Thr Asn His  
145 150 155 160

Gly Ser Leu Ser Asp Arg Ile Met Ser Tyr Tyr Gly Asp Ser Pro Arg  
165 170 175

Gly Met Val Glu Ser Ala Leu Glu Phe Ala Arg Ile Cys Arg Lys Leu  
180 185 190

Asp Phe His Asn Phe Val Phe Ser Met Lys Ala Ser Asn Pro Val Ile  
195 200 205

Met Val Gln Ala Tyr Arg Leu Leu Val Ala Glu Met Tyr Asn Leu Gly  
210 215 220

Trp Asp Tyr Pro Leu His Leu Gly Val Thr Glu Ala Gly Glu Gly Glu  
225 230 235 240

Asp Gly Arg Met Lys Ser Ala Ile Gly Ile Gly Thr Leu Leu Met Asp  
245 250 255

Gly Leu Gly Asp Thr Ile Arg Val Ser Leu Thr Glu Pro Pro Glu Glu  
260 265 270

Glu Ile Asp Pro Cys Arg Arg Leu Ala Asn Leu Gly Thr His Ala Ala  
275 280 285

Asp Leu Gln Ile Gly Val Ala Pro Phe Glu Glu Lys His Arg Arg Tyr  
290 295 300

Phe Asp Phe Gln Arg Arg Ser Gly Gln Leu Pro Leu Gln Lys Glu Ala  
305 310 315 320

Pro Glu Leu Leu Tyr Arg Ser Leu Ala Ala Lys Leu Val Val Gly Met  
325 330 335

Pro Phe Lys Asp Leu Ala Thr Val Asp Ser Ile Leu Leu Lys Glu Leu  
340 345 350

Pro Pro Val Glu Asp Ala Gln Ala Arg Leu Ala Leu Lys Arg Leu Val  
355 360 365

Asp Ile Ser Met Gly Val Leu Thr Pro Leu Ser Glu Gln Leu Thr Lys  
370 375 380

Pro Leu Pro His Ala Ile Ala Leu Val Asn Val Asp Glu Leu Ser Ser  
385 390 395 400

Gly Ala His Lys Leu Leu Pro Glu Gly Thr Arg Leu Ala Val Thr Leu  
405 410 415

Arg Gly Asp Glu Ser Tyr Glu Gln Leu Asp Leu Leu Lys Gly Val Asp  
420 425 430

Asp Ile Thr Met Leu Leu His Ser Val Pro Tyr Gly Glu Glu Lys Thr  
435 440 445



Gly Arg Val His Ala Ala Arg Arg Leu Phe Glu Tyr Leu Glu Thr Asn  
 450 455 460

Gly Leu Asn Phe Pro Val Ile His His Ile Glu Phe Pro Lys Ser Val  
 465 470 475 480

Asn Arg Asp Asp Leu Val Ile Gly Ala Gly Ala Asn Val Gly Ala Leu  
 485 490 495

Leu Val Asp Gly Leu Gly Asp Gly Val Leu Leu Glu Ala Ala Asp Gln  
 500 505 510

Glu Phe Glu Phe Leu Arg Asp Thr Ser Phe Asn Leu Leu Gln Gly Cys  
 515 520 525

Arg Met Arg Asn Thr Lys Thr Ile Ala Ile Met Gly Cys Ile Val Asn  
 530 535 540

Gly Pro Gly Glu Met Ala Asp Ala Asp Phe Gly Tyr Val Gly Gly Ala  
 545 550 555 560

Pro Gly Lys Ile Asp Leu Tyr Val Gly Lys Thr Val Val Gln Arg Gly  
 565 570 575

Ile Ala Met Glu Gly Ala Thr Asp Ala Leu Ile Gln Leu Ile Lys Asp  
 580 585 590

His Gly Arg Trp Val Asp Pro Pro Val Glu Glu  
 595 600

<210> 50  
 <211> 372  
 <212> PRT  
 <213> Escherichia coli  
 <400> 50

Met His Asn Gln Ala Pro Ile Gln Arg Arg Lys Ser Thr Arg Ile Tyr  
 1 5 10 15

Val Gly Asn Val Pro Ile Gly Asp Gly Ala Pro Ile Ala Val Gln Ser  
 20 25 30

Met Thr Asn Thr Arg Thr Thr Asp Val Glu Ala Thr Val Asn Gln Ile  
 35 40 45

Lys Ala Leu Glu Arg Val Gly Ala Asp Ile Val Arg Val Ser Val Pro  
 50 55 60

Thr Met Asp Ala Ala Glu Ala Phe Lys Leu Ile Lys Gln Gln Val Asn  
 65 70 75 80

Val Pro Leu Val Ala Asp Ile His Phe Asp Tyr Arg Ile Ala Leu Lys  
 85 90 95

Val Ala Glu Tyr Gly Val Asp Cys Leu Arg Ile Asn Pro Gly Asn Ile  
100 105 110

Gly Asn Glu Glu Arg Ile Arg Met Val Val Asp Cys Ala Arg Asp Lys  
115 120 125

Asn Ile Pro Ile Arg Ile Gly Val Asn Ala Gly Ser Leu Glu Lys Asp  
130 135 140

Leu Gln Glu Lys Tyr Gly Glu Pro Thr Pro Gln Ala Leu Leu Glu Ser  
145 150 155 160

Ala Met Arg His Val Asp His Leu Asp Arg Leu Asn Phe Asp Gln Phe  
165 170 175

Lys Val Ser Val Lys Ala Ser Asp Val Phe Leu Ala Val Glu Ser Tyr  
180 185 190

Arg Leu Leu Ala Lys Gln Ile Asp Gln Pro Leu His Leu Gly Ile Thr  
195 200 205

Glu Ala Gly Gly Ala Arg Ser Gly Ala Val Lys Ser Ala Ile Gly Leu  
210 215 220

Gly Leu Leu Leu Ser Glu Gly Ile Gly Asp Thr Leu Arg Val Ser Leu  
225 230 235 240

Ala Ala Asp Pro Val Glu Glu Ile Lys Val Gly Phe Asp Ile Leu Lys  
245 250 255

Ser Leu Arg Ile Arg Ser Arg Gly Ile Asn Phe Ile Ala Cys Pro Thr  
260 265 270

Cys Ser Arg Gln Glu Phe Asp Val Ile Gly Thr Val Asn Ala Leu Glu  
275 280 285

Gln Arg Leu Glu Asp Ile Ile Thr Pro Met Asp Val Ser Ile Ile Gly  
290 295 300

Cys Val Val Asn Gly Pro Gly Glu Ala Leu Val Ser Thr Leu Gly Val  
305 310 315 320

Thr Gly Gly Asn Lys Lys Ser Gly Leu Tyr Glu Asp Gly Val Arg Lys  
325 330 335

Asp Arg Leu Asp Asn Asn Asp Met Ile Asp Gln Leu Glu Ala Arg Ile  
340 345 350

Arg Ala Lys Ala Ser Gln Leu Asp Glu Ala Arg Arg Ile Asp Val Gln  
355 360 365

Gln Val Glu Lys  
370

<210> 51  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named CINCO

<400> 51

cgctgcccag aatggacctc cctag

25

<210> 52  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named SEIS

<400> 52

cagccgcggtt ttgacttgaa acgtgc

26

<210> 53  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named MPD-Nde5'

<400> 53

gccatatgac cgtttacaca gcatccg

27

<210> 54  
<211> 35  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named MPD-Eco3'

<400> 54

tcgaattctc attattcctt tggtagacca gtctt

35

<210> 55  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named hPMK1

<400> 55

tggttaacat atggccccgc tgggaggcgc

30

<210> 56  
<211> 40  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named hPMK4

<400> 56

aggttaactc aattaaagtc tggagcggat aaattctatc

40

<210> 57  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named UNO

<400> 57

cgggcctcgt ttggctgtcg cactg

25

<210> 58  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named DOS

<400> 58

cgcgggtgga aggaccttgt ggagg

25

<210> 59  
<211> 33  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named MK-Hpa5'

<400> 59

aagttaacat atgtcattac cggtcttaac ttc

33

<210> 60  
<211> 34  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named MK-Hpa3'

<400> 60

cggttaactc attatgaagt ccatggtaaa ttcg

34

<210> 61  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named idi5X

<400> 61

cccctcgaga ttatgcaaac ggaacacgtc

30

<210> 62  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named idi3X

<400> 62

ggctcgagtt atttaagctg ggtaaattgca g

31

<210> 63  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named pBAD-mut1

<400> 63

ctgagagtgc accatctgcg gtgtgaaata cc 32

<210> 64  
<211> 40  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named pBAD-Link1

<400> 64

aattctaagg aggttttaaac taaggaggta cgtaaggagg 40

<210> 65  
<211> 40  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named pBAD-Link2

<400> 65

tcgacctcct tacgtacctc cttagtttaa acctccttag 40

<210> 66  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named pBAD-D2

<400> 66

tcataactccc gccattcaga g 21

<210> 67  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named pBAD-U3

<400> 67

ccgccaaaac agccaagctt g 21

<210> 68  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named pRS-L1

<400> 68

gatccgttta aacgcccggg cggccgcg 28

<210> 69  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named pRS-L2

<400> 69

aattcgcggc cgcccgggcg tttaaacg 28

<210> 70  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named lPE

<400> 70

cgcggtgtgg gtgagcatga tg 22

<210> 71  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named 22PE

<400> 71

aaatctcccg ggttaccgt ctgttactgc

30

<210> 72  
<211> 33  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named 3PE

<400> 72

gcgttttaaac tggacgaagc gcgtcgaatt gac

33

<210> 73  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named 4PE

<400> 73

tgcacgaccg cccagttggt cc

22

<210> 74  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Designed primer named CAT1

<400> 74

gagtccgaat aaataacctgt g

21



<210> 75  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Designed primer named CAT4

<400> 75

ccgaatttct gccattcatc c 21

<210> 76  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Designed primer named OPE

<400> 76

tgggctttgt cacgagcaca c 21

<210> 77  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Designed primer named 5PE

<400> 77

ggcccatagc aaaaccgaca g 21

<210> 78  
 <211> 372  
 <212> PRT  
 <213> Escherichia coli

<400> 78

Met His Asn Gln Ala Pro Ile Gln Arg Arg Lys Ser Thr Arg Ile Tyr  
 1 5 10 15

Val Gly Asn Val Pro Ile Gly Asp Gly Ala Pro Ile Ala Val Gln Ser  
 20 25 30

Met Thr Asn Thr Arg Thr Thr Asp Val Glu Ala Thr Val Asn Gln Ile  
 35 40 45

Lys Ala Leu Glu Arg Val Gly Ala Asp Ile Val Arg Val Ser Val Pro  
 50 55 60

Thr Met Asp Ala Ala Glu Ala Phe Lys Leu Ile Lys Gln Gln Val Asn  
 65 70 75 80

Val Pro Leu Val Ala Asp Ile His Phe Asp Tyr Arg Ile Ala Leu Lys  
85 90 95

Val Ala Glu Tyr Gly Val Asp Cys Leu Arg Ile Asn Pro Gly Asn Ile  
100 105 110

Gly Asn Glu Glu Arg Ile Arg Met Val Val Asp Cys Ala Arg Asp Lys  
115 120 125

Asn Ile Pro Ile Arg Ile Gly Val Asn Ala Gly Ser Leu Glu Lys Asp  
130 135 140

Leu Gln Glu Lys Tyr Gly Glu Pro Thr Pro Gln Ala Leu Leu Glu Ser  
145 150 155 160

Ala Met Arg His Val Asp His Leu Asp Arg Leu Asn Phe Asp Gln Phe  
165 170 175

Lys Val Ser Val Lys Ala Ser Asp Val Phe Leu Ala Val Glu Ser Tyr  
180 185 190

Arg Leu Leu Ala Lys Gln Ile Asp Gln Pro Leu His Leu Gly Ile Thr  
195 200 205

Glu Ala Gly Gly Ala Arg Ser Gly Ala Val Lys Ser Ala Ile Gly Leu  
210 215 220

Gly Leu Leu Leu Ser Glu Gly Ile Gly Asp Thr Leu Arg Val Ser Leu  
225 230 235 240

Ala Ala Asp Pro Val Glu Glu Ile Lys Val Gly Phe Asp Ile Leu Lys  
245 250 255

Ser Leu Arg Ile Arg Ser Arg Gly Ile Asn Phe Ile Ala Cys Pro Thr  
260 265 270

Cys Ser Arg Gln Glu Phe Asp Val Ile Gly Thr Val Asn Ala Leu Glu  
275 280 285

Gln Arg Leu Glu Asp Ile Ile Thr Pro Met Asp Val Ser Ile Ile Gly  
290 295 300

Cys Val Val Asn Gly Pro Gly Glu Ala Leu Val Ser Thr Leu Gly Val  
305 310 315 320

Thr Gly Gly Asn Lys Lys Ser Gly Leu Tyr Glu Asp Gly Val Arg Lys  
325 330 335

Asp Arg Leu Asp Asn Asn Asp Met Ile Asp Gln Leu Glu Ala Arg Ile  
340 345 350

Arg Ala Lys Ala Ser Gln Leu Asp Glu Ala Arg Arg Ile Asp Val Gln  
355 360 365

Gln Val Glu Lys  
370

<210> 79  
 <211> 740  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 79

```

Met Ala Thr Gly Val Leu Pro Ala Pro Val Ser Gly Ile Lys Ile Pro
1          5          10          15

Asp Ser Lys Val Gly Phe Gly Lys Ser Met Asn Leu Val Arg Ile Cys
          20          25          30

Asp Val Arg Ser Leu Arg Ser Ala Arg Arg Arg Val Ser Val Ile Arg
          35          40          45

Asn Ser Asn Gln Gly Ser Asp Leu Ala Glu Leu Gln Pro Ala Ser Glu
          50          55          60

Gly Ser Pro Leu Leu Val Pro Arg Gln Lys Tyr Cys Glu Ser Leu His
65          70          75          80

Lys Thr Val Arg Arg Lys Thr Arg Thr Val Met Val Gly Asn Val Ala
          85          90          95

Leu Gly Ser Glu His Pro Ile Arg Ile Gln Thr Met Thr Thr Ser Asp
          100          105          110

Thr Lys Asp Ile Thr Gly Thr Val Asp Glu Val Met Arg Ile Ala Asp
          115          120          125

Lys Gly Ala Asp Ile Val Arg Ile Thr Val Gln Gly Lys Lys Glu Ala
          130          135          140

Asp Ala Cys Phe Glu Ile Lys Asp Lys Leu Val Gln Leu Asn Tyr Asn
145          150          155          160

Ile Pro Leu Val Ala Asp Ile His Phe Ala Pro Thr Val Ala Leu Arg
          165          170          175

Val Ala Glu Cys Phe Asp Lys Ile Arg Val Asn Pro Gly Asn Phe Ala
          180          185          190

Asp Arg Arg Ala Gln Phe Glu Thr Ile Asp Tyr Thr Glu Asp Glu Tyr
          195          200          205

Gln Lys Glu Leu Gln His Ile Glu Gln Val Phe Thr Pro Leu Val Glu
          210          215          220

Lys Cys Lys Lys Tyr Gly Arg Ala Met Arg Ile Gly Thr Asn His Gly
225          230          235          240

Ser Leu Ser Asp Arg Ile Met Ser Tyr Tyr Gly Asp Ser Pro Arg Gly
          245          250          255

```

Met Val Glu Ser Ala Phe Glu Phe Ala Arg Ile Cys Arg Lys Leu Asp  
 260 265 270  
 Tyr His Asn Phe Val Phe Ser Met Lys Ala Ser Asn Pro Val Ile Met  
 275 280 285  
 Val Gln Ala Tyr Arg Leu Leu Val Ala Glu Met Tyr Val His Gly Trp  
 290 295 300  
 Asp Tyr Pro Leu His Leu Gly Val Thr Glu Ala Gly Glu Gly Glu Asp  
 305 310 315 320  
 Gly Arg Met Lys Ser Ala Ile Gly Ile Gly Thr Leu Leu Gln Asp Gly  
 325 330 335  
 Leu Gly Asp Thr Ile Arg Val Ser Leu Thr Glu Pro Pro Glu Glu Glu  
 340 345 350  
 Ile Asp Pro Cys Arg Arg Leu Ala Asn Leu Gly Thr Lys Ala Ala Lys  
 355 360 365  
 Leu Gln Gln Gly Ala Pro Phe Glu Glu Lys His Arg His Tyr Phe Asp  
 370 375 380  
 Phe Gln Arg Arg Thr Gly Asp Leu Pro Val Gln Lys Glu Gly Glu Glu  
 385 390 395 400  
 Val Asp Tyr Arg Asn Val Leu His Arg Asp Gly Ser Val Leu Met Ser  
 405 410 415  
 Ile Ser Leu Asp Gln Leu Lys Ala Pro Glu Leu Leu Tyr Arg Ser Leu  
 420 425 430  
 Ala Thr Lys Leu Val Val Gly Met Pro Phe Lys Asp Leu Ala Thr Val  
 435 440 445  
 Asp Ser Ile Leu Leu Arg Glu Leu Pro Pro Val Asp Asp Gln Val Ala  
 450 455 460  
 Arg Leu Ala Leu Lys Arg Leu Ile Asp Val Ser Met Gly Val Ile Ala  
 465 470 475 480  
 Pro Leu Ser Glu Gln Leu Thr Lys Pro Leu Pro Asn Ala Met Val Leu  
 485 490 495  
 Val Asn Leu Lys Glu Leu Ser Gly Gly Ala Tyr Lys Leu Leu Pro Glu  
 500 505 510  
 Gly Thr Arg Leu Val Val Ser Leu Arg Gly Asp Glu Pro Tyr Glu Glu  
 515 520 525  
 Leu Glu Ile Leu Lys Asn Ile Asp Ala Thr Met Ile Leu His Asp Val  
 530 535 540  
 Pro Phe Thr Glu Asp Lys Val Ser Arg Val His Ala Ala Arg Arg Leu  
 545 550 555 560

Phe Glu Phe Leu Ser Glu Asn Ser Val Asn Phe Pro Val Ile His His  
565 570 575

Ile Asn Phe Pro Thr Gly Ile His Arg Asp Glu Leu Val Ile His Ala  
580 585 590

Gly Thr Tyr Ala Gly Gly Leu Leu Val Asp Gly Leu Gly Asp Gly Val  
595 600 605

Met Leu Glu Ala Pro Asp Gln Asp Phe Asp Phe Leu Arg Asn Thr Ser  
610 615 620

Phe Asn Leu Leu Gln Gly Cys Arg Met Arg Asn Thr Lys Thr Glu Tyr  
625 630 635 640

Val Ser Cys Pro Ser Cys Gly Arg Thr Leu Phe Asp Leu Gln Glu Ile  
645 650 655

Ser Ala Glu Ile Arg Glu Lys Thr Ser His Leu Pro Gly Val Ser Ile  
660 665 670

Ala Ile Met Gly Cys Ile Val Asn Gly Pro Gly Glu Met Ala Asp Ala  
675 680 685

Asp Phe Gly Tyr Val Gly Gly Ser Pro Gly Lys Ile Asp Leu Tyr Val  
690 695 700

Gly Lys Thr Val Val Lys Arg Gly Ile Ala Met Thr Glu Ala Thr Asp  
705 710 715 720

Ala Leu Ile Gly Leu Ile Lys Glu His Gly Arg Trp Val Asp Pro Pro  
725 730 735

Val Ala Asp Glu  
740

<210> 80

<211> 155

<212> DNA

<213> Arabidopsis thaliana

<400> 80

aaaaatcgga aaaatggcga ctggagtatt gccagctccg gtttctggga tcaagataacc 60  
ggattcgaaa gtcgggtttg gtaaaagcat gaatcttgtg agaatttgtg atgttaggag 120  
tctaagatct gctgatgagt agatttcata aaagt 155

<210> 81

<211> 42

<212> PRT

<213> Arabidopsis thaliana

<400> 81

Met Ala Thr Gly Val Leu Pro Ala Pro Val Ser Gly Ile Lys Ile Pro  
1 5 10 15

Asp Ser Lys Val Gly Phe Gly Lys Ser Met Asn Leu Val Arg Ile Cys  
 20 25 30

Asp Val Arg Ser Leu Arg Ser Ala Asp Glu  
 35 40

<210> 82  
 <211> 45  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 82

atgagaggat cgcaycayca ycaycaycay cayggatccg catgc

45

<210> 83  
 <211> 12  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 83

Met Arg Gly Ser His His His His His His Gly Ser  
 1 5 10

<210> 84  
 <211> 59  
 <212> DNA  
 <213> Arabidopsis thaliana

<400> 84

atgagaggat cgcaycayca ycaycaycay ggatctgctg atgagtagat ttcgcatgc

59

<210> 85  
 <211> 15  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 85

Met Arg Gly Ser His His His His His His Gly Ser Ala Asp Glu  
 1 5 10 15